New materials for micro-scale sensors and actuators
An engineering review

Stephen A. Wilson a,*, Renaud P.J. Jourdain a, Qi Zhang a, Robert A. Dorey a, Chris R. Bowen b,1, Magnus Willander c,2, Qamar Ul Wahab d,3, Magnus Willander e,4, Safaa M. Al-hilli e,4, Omer Nur e,4, Eckhard Quandt f,5, Christer Johansson g,6, Emmanouel Pagounis h,7, Manfred Kohl i,8, Jovan Matovic j,9, Björn Samel k,10, Wouter van der Wijngaart k,10, Edwin W.H. Jager l,11, Daniel Carlsson l,11, Zoran Djinovic j,12, Michael Wegener p,13, Carmen Moldovan m,14, Rodica Iosub m, Estefania Abad n,15, Michael Wendlandt o,16, Cristina Rusu g,17, Katrin Persson g,17

a Microsystems and Nanotechnology Group, Materials Department, Cranfield University, Cranfield, Bedfordshire MK43 0AL, United Kingdom
b Materials Research Centre, Department of Mechanical Engineering, University of Bath, Bath BA2 7AY, United Kingdom
c Physical Electronics, Department of Science and Technology, Campus Norrköping, Linköping University, SE-601 74 Norrköping, Sweden
d Department of Physics, Measurement Technology, Biology and Chemistry, Linköping University, SE-581 83 Linköping, Sweden
e Physical Electronics and Photonics, Physics Department, Gothenburg University, SE-412 96 Gothenburg, Sweden
f Institute for Materials Science, Faculty of Engineering, University Kiel, Kaiserstr. 2, 24143 Kiel, Germany
g Imego AB, Arvid Hedvalls Backe 4, SE-411 33 Göteborg, Sweden
h Helsinki University of Technology, Laboratory of Materials Science, Vuorimiehentie 2A, 02015 TKK, Finland
i Microsystems, Forschungszentrum Karlsruhe, IMT, Postfach 3640, 76021 Karlsruhe, Germany
j Institute of Sensor and Actuator Systems, Vienna University of Technology, Floragasse 7/2, A-1040 Vienna, Austria
k Microsystems Technology Lab (MST), School of Electrical Engineering (EE), Royal Institute of Technology (KTH), Osquldas vag 10, S-100 44 Stockholm, Sweden
l Micromuscle AB, Teknikringen 10, SE-583 30 Linköping, Sweden
m Microstructures for Bio-Medical Applications Research Laboratory, National Institute for Research and Development in Microtechnologies, IMT-Bucharest, 31B Erou Iancu Nicolae Street, 077190 Bucharest, Romania
n Micro and Nanotechnology Department, Fundación Tekniker, Avenida Otuola 20, 20600 EIBAR (Guipuzcoa), Spain
o Micro and Nanosystems, Department of Mechanical Engineering, ETH Zurich, 8092 Zurich, Switzerland
p Functional Polymer Systems, Fraunhofer Institute for Applied Polymer Research, Geiselbergstrasse 69, 14476 Potsdam-Golm, Germany

Received 23 February 2007; received in revised form 20 March 2007; accepted 20 March 2007
Available online 29 June 2007
Abstract

This paper provides a detailed overview of developments in transducer materials technology relating to their current and future applications in micro-scale devices. Recent advances in piezoelectric, magnetostrictive and shape-memory alloy systems are discussed and emerging transducer materials such as magnetic nanoparticles, expandable micro-spheres and conductive polymers are introduced. Materials properties, transducer mechanisms and end applications are described and the potential for integration of the materials with ancillary systems components is viewed as an essential consideration. The review concludes with a short discussion of structural polymers that are extending the range of micro-fabrication techniques available to designers and production engineers beyond the limitations of silicon fabrication technology.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Piezoelectric; Magnetic; Shape memory; Polymer; Microstructure; Microtechnology

Contents

1. Introduction ................................................................. 5
2. Ferroelectric ceramics ..................................................... 6
   2.1. Piezoelectric properties and potential applications of ferroelectric thin films ......... 7
   2.1.1. Thin film deposition ..................................................... 8
   2.1.2. Piezoelectric properties of ferroelectric thin films ....................... 8
   2.1.3. Poling and reliability issues ........................................... 9
   2.1.4. Summary—ferroelectric thin films .................................... 10
   2.2. Thick film fabrication for micro-scale sensors ............................... 10
   2.2.1. Thick film deposition techniques ................................... 10
   2.2.2. Inks ..................................................................... 11
   2.2.3. Transformation binders ................................................. 12
   2.2.4. Electrical properties of PZT thick films ............................... 12
   2.2.5. Summary—ferroelectric thick films .................................. 12
3. Piezoelectric semiconductors ............................................. 13
   3.1. Groups III–V nitrides (GaN/AlN) .......................................... 13
   3.2. Groups III–V materials .................................................. 15
   3.3. ZnO materials .......................................................... 15
   3.4. Summary—piezoelectric semi-conductors ................................. 16
4. Zinc oxide structures for chemical sensors ............................. 16
   4.1. Synthesis and properties of ZnO nano-structures ......................... 17
   4.2. Electrochemical potential method ........................................ 18
   4.3. Site binding method ...................................................... 19

1 Tel.: +44 1225 383660; fax: +44 1225 826098.
2 Tel.: +44 11 363167.
3 Tel.: +46 13 288936.
4 Tel.: +46 31 772 2093/2097/3487; fax: +46 31 772 2092.
5 Tel.: +49 431 880 6200; fax: +49 431 880 6203.
6 Tel.: +46 31 750 18 61; fax: +46 31 750 18 01.
7 Tel.: +358 40548321; fax: +358 94512677.
8 Tel.: +49 7247 82x2798; fax: +49 7247 827798.
9 Tel.: +43 2622 22859 21; fax: +43 2622 22859 17.
10 Tel.: +46 8 790 66 13; fax: +46 8 10 08 58.
11 Tel.: +46 13 3420053; fax: +46 13 3420059.
12 Tel.: +43 2622 22859 21; fax: +43 2622 22859 17.
13 Tel.: +49 331 568 1209; fax: +49 331 568 3910.
14 Tel.: +40 21 4908212; fax: +40 21 49082381.
15 Tel.: +34 943 20 67 44; fax: +34 943 20 27 57.
16 Tel.: +41 632 47 05; fax: +41 632 14 62.
17 Tel.: +46 31 75018 68; fax: +46 31 75018 01.
14.3. Piezoelectric properties ............................................................ 81
14.4. Applications of ferroelectrets ........................................................ 82
14.5. Conclusions and outlook ............................................................. 82
15. Conductive polymers ................................................................. 83
15.1. Mechanism of polymer conductivity—role of doping ......................................... 83
15.2. Conductive polymeric materials—examples ..................................................... 85
15.2.1. Polypyrrole .............................................................................. 85
15.2.2. Polyaniline .............................................................................. 85
15.2.3. Polylithiophene ...................................................................... 85
15.2.4. Polysiloxane ........................................................................... 86
15.2.5. Polylithalocyanine .................................................................. 86
15.2.6. Fullerene ................................................................................. 87
15.3. Applications of conductive polymers in sensors and actuators ................................ 87
15.3.1. Sensors .................................................................................... 87
15.3.2. Chemical microsensors .............................................................. 88
15.3.3. Electronic noses ...................................................................... 89
15.3.4. FET type devices .................................................................... 90
15.3.5. Biosensors ............................................................................... 91
15.3.6. Actuators ................................................................................ 91
15.4. Conclusions .............................................................................. 92
16. Polyimides .............................................................................. 93
16.1. Properties of polyimides ...................................................................... 93
16.2. Processing of polyimides.............................................................................. 93
16.2.1. Wet etch patterning .................................................................. 93
16.2.2. Dry etch patterning .................................................................. 94
16.2.3. Photodefined polyimides ............................................................ 94
16.2.4. Laser ablation ......................................................................... 95
16.3. Polyimide applications .................................................................... 95
16.3.1. High density interconnection flexible substrates ..................................... 95
16.3.2. MEMS devices ....................................................................... 95
17. Structural polymers ...................................................................... 97
17.1. Selection of structural polymers for micro-scale devices .................................... 98
17.1.1. Thermosets ............................................................................ 98
17.1.2. Thermoplastics ....................................................................... 100
17.1.3. Elastomers ............................................................................ 101
17.2. Applications ............................................................................... 101
17.2.1. Micro-scale sensors ................................................................. 101
17.2.2. Micro-scale actuators ............................................................... 102
18. Integration and interconnection .......................................................... 103
18.1. Wafer bonding ............................................................................ 104
18.1.1. Adhesive bonding .................................................................... 104
18.1.2. Metallic bonding ..................................................................... 105
18.1.3. Glass-frit bonding ................................................................... 105
18.1.4. Silicon direct bonding ............................................................. 105
18.1.5. Plasma-enhanced bonding ....................................................... 106
18.1.6. Anodic bonding ..................................................................... 106
18.2. Low temperature co-fired ceramics and microsystems ...................................... 107
18.2.1. Medium CTE LTCC .............................................................. 108
18.2.2. Low CTE LTCC ................................................................... 109
18.3. Characterisation methods for microsystem bonding ......................................... 110
18.4. Conclusion ................................................................................ 112
Acknowledgements ............................................................................ 112
References ....................................................................................... 112
1. Introduction

A material can be said to be ‘new’ or ‘novel’ until it finds its way into mainstream engineering technology. The distinguishing criterion is not whether the end-use is in consumer products, sophisticated, specialised or niche applications, but whether materials performance is predictable and reliable. By implication, quality and processing must be well understood and commercial supplies readily available. For these reasons, the time-scale in which a material remains new is related directly to the commercial interest that has evolved and consequently to the business opportunities that the material has inspired in its conceptual form.

A new material that promises to provide tangible improvements over the established norm will soon attract commercial interest and its’ potential use will come under scrutiny. The first questions to be raised relate to possible integration into existing systems or possible creation of a new product line. If technological barriers to integration exist, be these either real or perceived, then commercial interest will immediately cool. For the particular case of micro-systems technology (MST), where the creation of fine scale integrated systems is a key motivational factor, the potential costs of product development can often overshadow any improvements in performance that might be gained. This is partly a consequence of local integration with microelectronics and packaging and it is partly due to the capital equipment costs involved. In the main, however, it is due to the time and uncertainty involved in establishing a new fabrication route that meets predefined standards of quality and reliability. Hence, to gain acceptance in micro-technology the new material must offer distinct performance advantages and it must also be compatible with various ancillary systems components and packaging. In all cases, it is highly probable that production will entail a lengthy sequence of process steps and consequently the material will need to tolerate repeated thermal cycling as system fabrication proceeds. It is not uncommon for the materials covered in this review, namely transducer materials, to rely on some aspect of their micro-structural composition that is highly sensitive to processing conditions. As an example, effects of grain size or morphology are often critical and optimum performance can be impaired by excursions outside a limited temperature range. Therefore, the processes involved in creating the material may only be one part of the equation and compatibility with secondary systems fabrication processes is equally essential.

Full-integration of micro-electronic and micro-mechanical components on a single wafer has been achieved commercially using silicon processing technology. Some examples of products made in this way include micro-gyroscopes and micro-mirror arrays. Whilst this integrated design approach appears to be commercially attractive it has, however, proven to be relatively rare owing to the complexity of the design process and, consequently, high development costs. Furthermore, due to processing restrictions the mechanical components of these fully-integrated devices are often constructed simply from silicon and silicon oxide with selective metallization. An alternative approach, adopted much more commonly, is via a hybrid design where component parts are created separately for subsequent assembly into a complete system. For small or medium-scale batch production this is an attractive option, as it removes many of the restrictions imposed by the need for process compatibility. Furthermore, test procedures can be performed at the wafer-scale before final assembly to enhance quality and overall yield. It is in this context that new transducer materials have the best chance of success. Key considerations are the availability of material-specific replication technologies, device-specific geometric requirements (feature types, planar or 3D, aspect ratios), the required dimensional tolerances and accuracy, surface quality or integrity, volumetric production rate and material cost, which can often be of secondary importance in this context.

Overall it can be said that the most significant barriers to progress are firstly the availability of production technologies and secondly the availability of knowledge. This article therefore seeks to review recent developments in transducer materials technology and to place them in the context of their current and future applications in micro-scale systems fabrication. In addition to examining recent advances in piezoelectric, magnetostrictive and shape memory alloys systems, emerging transducer materials such as magnetic nanoparticles, expandable micro spheres and conductive polymers are also discussed. Their underlying properties, transducer mechanism and end applications are described, along with the processing technologies to form them in particulate, bulk or film geometry. Aspects of processing that may influence integration of the materials with their related components are viewed as an essential consideration. From a global perspective, there are of necessity some important omissions. It seems certain that materials incorporating carbon nanotube technology and nanocomposites will reach industrial maturity in the very near future and that their impact will be significant. This subject matter has been extensively reviewed elsewhere and
the materials are not covered in this review. Rather the intention is to highlight a range of materials that could be used in conjunction with standard micro-fabrication techniques to extend the range of devices that can be made beyond the limitations of silicon fabrication technology.

2. Ferroelectric ceramics

Polycrystalline lead zirconate titanate (PZT) ceramics are of major importance in microtechnology, particularly in the field of sensors and actuators, because of their superior piezoelectric and pyroelectric properties and their high dielectric constants \[1\]. Devices that incorporate these materials as their active component include micro-pumps and valves, ultrasonic motors, thermal sensors, probes for medical imaging and non-destructive testing, accelerometers and quite recently a new range of electronic components that includes filters, memory devices and switches. New applications continue to emerge, and a major research effort has been underway to address the manufacturing technology required to incorporate these materials with associated structural components and electronic circuitry at the wafer scale. Two distinct approaches are available, which have very different process requirements and which consequently require different fabrication techniques. The bottom-up approach is by thin film deposition, performed via spin coating of a sol–gel precursor or sputtering. Thin film compositions have been developed that have greatly reduced processing temperatures (600–700 °C) in comparison to standard bulk ceramic sintering (1100–1400 °C) and this has led to commercialization by the major electronics corporations in the form of ferroelectric memories and electronic components. A single layer is typically around 0.1 mm and films are built up to the required thickness by depositing several layers in succession.

The processing issues that surround production of electromechanical devices on the micro-scale are arguably even more complex, however, due to the range of ancillary system components that are needed. The available force that can be generated by the ceramic is directly related to the amount of electro-active material that is available and many piezoelectric devices with potential commercial applications such as micro-pumps require much thicker films to be effective, typically in the size range 10–80 μm. These values have been achieved by multi-layer deposition using composite thick film techniques and significant progress has been made, which makes these materials suitable for a number of applications. This technique is detailed below. In practice residual tensile stress is a critical issue, inherent to the process, which becomes progressively more significant as film thickness increases. Tensile stresses result from substrate clamping as the material crystallizes at elevated temperatures often leading to reduced fracture toughness or cracking and somewhat lower electro-active coefficients.

The alternative, top-down approach for micro-scale device fabrication is by assembly of net shape components, usually by adhesive bonding. This is routinely adopted for one-off device fabrication in the research environment. On the wafer scale, there are important questions of positional accuracy both laterally and in terms of parallelism with underlying materials. This becomes more significant as layer thicknesses are reduced below ~80 μm. The nature of the bond is of critical importance to device performance and hence the surface roughness and particularly the flatness of the ceramic component are very significant. Recently, it has been shown that bulk PZT ceramics can be thinned in situ to thicknesses well below 50 μm, using ultra-precision grinding, after bonding to wafer-scale components \[2\]. This technique has several advantages: (a) the electro-active properties of the ceramic can be fully exploited; (b) films can be made in the 20–50 μm thickness range, which is difficult to achieve by other methods; (c) ceramic films can be engineered into residual compression to optimize device performance; (d) the machining techniques can be used in sequence with standard micro-fabrication processes, such as photolithography, without the need for a high temperature excursion, thereby extending design flexibility and the range of devices that can be produced; (e) PZT films in this thickness range can be activated well below 100 V, this is highly significant in commercial terms as they are then compatible with current CMOS drive circuitry. Recent research work in this area has led to major improvements in technique and the method can be considered viable for flexible, batch-scale assembly and systems integration. The key issues that are involved in producing exceptionally smooth, flat surfaces in PZT by means of ultra-precision grinding have been discussed by Arai et al. \[3–5\].

As noted, ferroelectric ceramics are of widespread technological importance and for this reason they remain the subject of intense research activity. Materials development has focussed on three particular areas. One of these can be...
said to be market-driven through strong commercial interest in new fuel injection systems for motor vehicles. This is a high power, high temperature, low voltage application which is satisfied by multi-layer ceramic stacks. The ceramic layers are typically less than 50 μm in thickness and they are co-fired with metallic interlayers to produce an inter-digitated structure. As the layers are thin a low applied voltage can be used to generate a strong electric field in the ceramic [4]. A further area of both commercial and technological interest is in high frequency medical ultrasonics for imaging and ultrasound-guided therapy. This also tends to be a high power application where the goal is to reduce the energy losses that result from internal power dissipation. These can generate significant amounts of heat leading to thermal instability and loss of performance [6–8].

The second major area of research is pushed by new technology that has emerged in the form of ferroelectric single crystal materials. This type of material has recently become available in commercial quantities and the electro-active properties exhibited are a marked extension beyond those of conventional polycrystalline ceramics. The crystals are relaxor ferroelectric materials and they are typically based on the lead magnesium niobate–lead titanate (PMN-PT) solid solution, although many other compositions are also in research. Relaxors are characterised by a diffuse dielectric phase transition, that is to say their dielectric permittivity is both frequency and temperature dependent. Their physical behaviour is as yet not fully understood but, importantly, they are found to exhibit very large dielectric permittivities and very high piezoelectric coefficients. In operation, their electro-mechanical behaviour is predominantly electrostrictive in nature resulting in exceptionally low hysteretic losses even at high frequencies. Whilst these materials have shown clear superiority for some electro-acoustic applications, their adoption for use in actuators is still at a very early stage. The upper temperature limit of operation can be relatively low at around 50–80 °C and this, together with a marked environmental variability of properties, clearly imposes some restrictions on design. Nevertheless, these materials do show very interesting new capabilities and they are an exciting technological innovation [9–15].

The third main focus of research is driven by environmental concerns over the industrial use of compounds containing lead. Whilst it can be argued that the toxicity of lead-containing ceramics or glasses is very significantly reduced in comparison to that of the base metal, there is pressure to reduce its consumption. This has led to a concerted effort world-wide to identify equivalent electro-active materials that are lead-free. To-date, despite some significant investment of time and resources, little progress has been made in developing materials that are able to outperform standard PZT ceramics. Several interesting compositions have been identified, however, that have useful transducer properties and work seems sure to continue [16–20].

2.1. Piezoelectric properties and potential applications of ferroelectric thin films

Thin films are generally considered to have thicknesses less than 1 micron. Interest in ferroelectric thin films has been considerable over the last 20 years, driven by the possibility of using them for non-volatile memory applications and new microelectromechanical systems (MEMS). Thin film piezoelectric materials also offer a number of advantages in MEMS applications, due to the relatively large displacements that can be generated, the high energy densities, as well as high sensitivity sensors with wide dynamic ranges and low power requirements [21].

Piezoelectric MEMS devices contain at least two elements: a bulk silicon frame and a piezoelectric deflection element built onto it, which also has electrodes to apply or detect voltage potentials. The silicon substrate often provides only the structural element, defining the mechanical properties, while the added functional material such as piezoelectric thin films provide a direct transformation between a driving signal or a read-out signal and a sensor or an actuator parameter.

A sampling of recent developments in piezoelectric transduction devices using thin films includes lead zirconate titanate (PZT) based ultrasonic micromotors [22–24], cantilever actuators, probes for atomic force microscopy [25], micropumps [26], ultrasonic transducers for medical applications [27,28] and uncooled thermal imaging as pyroelectric arrays [29,30]. The aims of this section are as follows:

- To introduce the current fabrication techniques for piezoelectric thin films.
- To discuss the important piezoelectric coefficients and the key issues or factors influencing the piezoelectric properties of ferroelectric thin films.
- To discuss piezoelectric thin film poling and reliability issues.
2.1.1. Thin film deposition

Most of the existing physical and chemical coating techniques have been investigated for the deposition of PZT. The physical methods include ion beam sputtering [31], rf magnetron sputtering [32,33], dc magnetron sputtering [34] and pulsed laser deposition (PLD) [35–37]. Chemical methods include metal-organic chemical vapour deposition (MOCVD) [38–42] and chemical solution deposition (CSD) [43,44]. Today there is a clear trend to apply MOCVD or CSD since a particular advantage with MOCVD is that conformal coating of three-dimensional objects is possible. CSD is a low cost technique for small-scale production, as required in the sensor industry. Since for CSD the film is initially amorphous, post-annealing treatments are necessary to crystallize the film. All the other methods described above allow in situ growth. Although the CSD technique seems very different from the vacuum deposition techniques like sputtering or PLD, there are nevertheless some common features:

- The crystallinity and texture of the film are strongly dependent on the crystal structure of the substrate, for example: lattice parameters and thermal expansion coefficients matching, surface defects, etc.
- The quality of the interface is dependent on the substrate chemistry, for example: reactivity of the substrate surface with the deposited phase constituents, diffusion coefficients, etc.
- The lattice energy has to be brought to the system, either thermally or by a physical way, since the initial state is a disordered one (gas or liquid phase, plasma, particle beam, etc.).
- Nucleation and growth of the perovskite require a precise stoichiometry, otherwise competing phases with fluorite (Pb2+xTi2-xO7) and pyrochlore (PbTi3O7) structures will nucleate [45].
- The growth is nucleation controlled [46,47].

2.1.2. Piezoelectric properties of ferroelectric thin films

The piezoelectric properties of ferroelectric materials, such as PbZr1-xTi xO3, are highly dependent on composition [21]. A schematic diagram of the lead zirconate (PZ)–lead titanate (PT) phase diagram is shown in Fig. 1. PZT has two main ferroelectric phases; rhombohedral for \( x < 0.48 \) and tetragonal for \( x > 0.48 \) under standard conditions. The rhombohedral phase is divided into ‘high temperature’ and ‘low temperature’ phases with crystal symmetries R2m and R3c, respectively. The boundary between the tetragonal and rhombohedral phases is sharply defined and virtually independent of temperature and the boundary is known as the morphotropic phase boundary (MPB). The boundary was defined by Jaffe et al. [48] to be at a composition of 53 % Zr and 47% Ti in PZT ceramics, and is defined as the point of equal coexistence for tetragonal/rhombohedral phases. In bulk ceramics, maxima in the piezoelectric coefficients are generally observed at the MPB. The same behaviour is often [49–55], but not universally [54–56], reported in thin films.

In MEMS technology, most of the piezoelectric thin films are polycrystalline materials. The piezoelectric effect is averaged over all the grains. The optimum piezoelectric properties of ferroelectric materials can only be obtained for

![Fig. 1. Phase diagram of the PbZrO3–PbTiO3 system [48].](image-url)
polycrystalline materials after an appropriate ‘poling’ treatment. Poling is the term used to describe a preliminary procedure that must be carried out, whereby a strong electric field is used to switch the initial, quasi-random internal polarisation of the poly-domain structure into a meta-stable alignment in the direction of the applied field. As a result, there is a net polarisation and a net piezoelectric effect. This can simplify processing, since single crystals are not required for good electromechanical properties.

The piezoelectric properties of films are almost always smaller than those of corresponding bulk ceramics. This is due to substrate clamping, which reduces the amount of strain which the film can exhibit for a given applied electric field or stress [56,57]. The film is part of a composite structure consisting of the piezoelectric film and silicon substrate. The film is effectively clamped in the film plane, but free to move in the out-of-plane direction. Therefore, the clamping effect is thickness dependent, and the piezoelectric coefficients, such as \( d_{33,f} \), increase with increasing thickness over a range of film thickness [21,58–62]. In thin film ceramics, it is conventional to assign the index 3 to the poling direction, usually perpendicular to the film plane. The directions of 1, 2 are therefore in the plane of the film. In a polycrystalline film, directions 1 and 2 are equivalent which implies that the in-plane strains (\( d_{31} \) and \( d_{32} \)) due to an applied electric field though the film thickness (\( E_3 \)) are isotropic and \( d_{31} = d_{32} \).

The relative coefficients of piezoelectric thin films are the effective values of \( d_{33,f} \) and \( e_{31,f} \), which are obtained as follows from the bulk tensor properties [63,56]:

\[
\begin{align*}
\frac{d_{33,f}}{d_{33}} &= \frac{1}{\left(\frac{s_{11}}{s_{11}} + \frac{s_{12}}{s_{12}}\right)} \\
\frac{e_{31,f}}{e_{31}} &= \frac{d_{31}}{\left(\frac{s_{11}}{s_{11}} + \frac{s_{12}}{s_{12}}\right)} 
\end{align*}
\]

The \( d_{33,f} \) coefficient can be directly measured as the strain per unit electric field through the film thickness (\( x_3/E_3 \)) provided that \( x_1 = x_2 = \sigma_3 = 0 \), where \( x_1 \) and \( x_2 \) are in-plane strains, \( \sigma_3 \) off-plane stress, \( x_3 \) is off-plane strain and \( s_{ij} \) is a compliance of the thin film. This measurement has been achieved with a double-beam Mach-Zehnder interferometer [64] that measures the thickness change of a film clamped on a much thicker substrate (assuring \( x_1 = x_2 = 0 \) at \( \sigma_3 = 0 \). The measurement of the transverse piezoelectric coefficient \( e_{31,f} \) has been undertaken with a cantilever bending method, collecting the charges as a function of \( x_1 \) and \( x_2 \) at zero \( \sigma_3 \) and electric field [65].

Apart from mechanical clamping due to the inert substrate, there are several other factors which influence the piezoelectric response of ferroelectric thin films, including orientation of the film [50,66–68], grain size [69], the level of polarization and breakdown field strength [70,71]. The influence of defects on the domain-wall contributions to the piezoelectric effect in thin films has not yet been studied in detail. Thus, it is presently not clear whether, for example, the effect of acceptor and donor dopants on the properties of PZT films would lead to the same effects as in bulk materials.

Film orientation can have a substantial effect on piezoelectric coefficients. Piezoelectric coefficients are optimized when the polarization axis, namely \( c \)-axis or \((0 \ 0 \ 1)\), is perpendicular to the film surface. It has been recently demonstrated [58] that the sol–gel derived PZT thin films with higher \( c \)-axis orientation exhibited larger piezoelectric coefficients. For random polycrystalline films, poling is often necessary to reorient the domains along the poling direction.

In many of the structures applied to MEMS technology, the piezoelectric film is part of a composite structure, i.e. the piezoelectric film is clamped to another elastic body. The coupling coefficient not only depends on the material parameters, but film stresses also play a role and such film stresses introduced during processing at elevated temperature are unavoidable. The residual stress can be as high as 10–100 MPa [72], which induces a pre-strain, or a pre-curvature to micromechanical structures. This stress has to be taken into account in the design phase of the devices.

2.1.3. Poling and reliability issues

The effects of poling in thin films differ from that in ceramics, since the clamping effect of the substrate pins the motion of \( a \)-domains [56,73]. In bulk ceramics, the clamping is effectively zero, and domains are relatively free to move in alignment with the poling field. There are few studies to date that are specifically related to thin film poling for
piezoelectric measurement, but it is well known that the strain induced by poling can be close enough to the tensile strength of the film which can induce cracking or delamination. Poling usually takes place at elevated temperatures (<150 °C) and at high field (200–300 kV/cm) as this increases domain wall mobility and enables better alignment along the field direction. Some examples of PZT thin film devices are shown in Fig. 2.

A further important point of performance is stability during operation and with time. The effective measured piezoelectric coefficients decay with time after poling in a process known as piezoelectric ageing, during which the domains in the poled sample revert to a more thermodynamically stable configuration. Depolarization (fatigue) may occur and, if integration of the film into the MEMS structure is not optimised, delamination of the PZT film or the electrodes may occur [74]. From an industrial point of view, the evaluation of ageing and fatigue is certainly an important task, however, only limited studies have been reported so far [75–77].

2.1.4. Summary—ferroelectric thin films

Ferroelectric thin films continue to represent an area of dynamism and technical advance in MEMS. Over the last 20 years, considerable progress has been made in optimizing the deposition conditions for thin films to improve the available piezoelectric activity although the growth of good quality PZT thin films still requires some effort. In processing such films, wet chemical methods continue to appear attractive for many applications. Recently, the attention has shifted from preparing novel ferroelectric films to the integration of such films in complex devices.

The overall estimation of performance is best seen in device applications since the performance of the devices depends not only on the properties of the materials, such as film orientation, grain size, thickness, etc., but also the composite structure of the devices in many cases.

In the future, the materials community requires greater knowledge of, and ability to control, the microstructure of films, and much more effective interaction with device technologists to bring commercial systems into widespread use.

2.2. Thick film fabrication for micro-scale sensors

Thick films are generally considered to be those with thicknesses greater than 1 μm, however, such a definition is imprecise as many thin film technologies can now achieve film thicknesses in excess of 1 μm. Thick films are required to increase the amount of functional material present in order to achieve higher displacements or increased power compared to thin films, e.g. for acoustic transducers or micro pumps (Fig. 3). For the purposes of this discussion, PZT thick films will be considered to be those that are formed using a powder suspension based processing route. These suspensions are typically made up of the desired ceramic powder (to impart the required functional properties), a carrier fluid and additives designed to improve the stability of the ink and processing of the ceramic material. For further information on issues associated with thick film processing and patterning of thick film structures the reader is directed towards an earlier review [78].

2.2.1. Thick film deposition techniques

Many different forming techniques can be used to deposit thick films due to the ability to tailor the fluidic characteristics (e.g. surface tension, viscosity, shear behaviour) of the powder suspensions. Despite the difference in
processing techniques the general processing stages are retained. Firstly, a suspension of the powder is deposited onto a substrate. Drying of the suspension then causes solvents and other additives to be removed. Finally, a high temperature stage is used to sinter the film.

_Screen printing_ is the most widely used thick film deposition technique [79] due to the ability to simultaneously create the thick film and pattern it. During the printing process, ink (containing the powder suspension) is forced through a fine mesh to deposit it onto the desired substrate. The mesh can be selectively masked off to enable a desired pattern to be created. The rheology of the ink is such that it does not pass through the mesh when at rest. When a shear stress is applied by the ‘squeegee’ of the screen printing device the viscosity decreases by shear thinning and the ink passes through the mesh. The ultimate resolution of the screen printing process is limited by the resolution of the mesh and the flow of the ink once printed.

Once the ink has been printed it is dried to remove the solvents. Subsequent layers can then be deposited prior to removing the organic components, such as polymers and modifiers, at temperatures between 350 and 600 °C. Final sintering occurs at temperatures between 850 and 950 °C.

Inks with lower powder loadings and viscosities (and less shear thinning behaviour) are used with processes such as _spin coating_ [80], _dip coating_ [81] and _spray coating_ [82,83]. All of these techniques result in the formation of continuous films. By further reducing the powder loading (typically below 1 vol% [84]) _electrophoretic deposition_ (EPD) can be used to create continuous films. In EPD a DC electric field (either constant field or constant current density) is used to attract charged ceramic particles to a substrate [85,86] with the advantage that complex geometries can be coated [87]. The limitation of the EPD process is that the substrate must be conducting, which may present difficulties in MEMS devices, and high density systems (e.g. lead based) are difficult to stabilise [85]. A thorough review of the electrophoretic deposition technique, as applied to many ceramic oxide films, is given by Sarkar and Nicholson [88].

### 2.2.2. Inks

In addition to using different deposition routes, it is also possible to use inks based on different principles—non-fugitive binders, fugitive binders and transformation binders. The most straightforward ink type is the _non-fugitive binder_ type where the binder remains within the system and forms the structural component of the film. Paints typify such systems, where an organic binder is used to bind the film together and bond the film to the substrate. PZT paints have been used to construct strain sensors on large structures [89]. The advantage of such systems is that no (or minimal) heating is required for processing the films. This means that the films can be applied to delicate substrates or large structures where it would be impractical to apply a post heat treatment. The disadvantage is that the low percentage (often little more than 50%) and low interconnectivity of active material within the films means that the functional properties of the films are significantly below those of bulk materials and conventionally sintered thick films.

The majority of ceramic powder processing is based on _fugitive binder_ systems where a temporary binder is used to impart limited strength to the system while it is shaped and handled. Subsequent thermal processing is then used to remove the binder and cause sintering of the ceramic phase. These fugitive binder inks use the same principle with inks containing the powder, and organic binder, a carrier fluid, and additives. The carrier fluid (usually water or a solvent) allows the powder to be handled conveniently and shaped. Once the film has been deposited and the carrier fluid
removed through evaporation, the film is held together and to the substrate by the fugitive binder phase. To enhance the sintering kinetics of the system, and lower the sintering temperature, a sintering aid is often added to the ink. These sintering aids form a liquid at temperatures in the region of 700–900 °C, which initially facilitates the reorientation of the ceramic particles to enhance the packing. Once reorientation has occurred the liquid phase sintering aid also acts as a fast diffusion path for atomic species and so encouraging the densification process to occur.

The sintering temperature of bulk PZT is typically between 1100 and 1300 °C. With liquid phase sintering aids these temperatures can be lowered to between 800 and 900 °C. Examples of sintering aids used for lead based piezoelectric materials include PbO [90–93], PbO–CuO [80,94,95], Pb5Ge3O11 [96–98], LiBiO2–CuO [99], PbO–PbF2 [100–102], Bi2O3–B2O3–CdO [100,103,104], borosilicate glass [103,105,106,93], Li/PbO [85] and PbO/TiO2 [107]. Although sintering aids significantly reduce the sintering temperature and enhance densification the presence of significant levels of non/low-function material within the thick film can reduce the resultant functional and piezoelectric properties. Furthermore, as sintering aids enhance the degree of solution-reprecipitation of the major phase (i.e. PZT) there is a large degree of atomic mixing. This may affect the electromechanical properties of the ceramic, i.e. if a soft doped PZT is the major phase, but the sintering aid contains a significant proportion of hard dopants then the sintered PZT material will behave as a hard doped material. Care should therefore be taken when selecting the appropriate sintering aid.

2.2.3. Transformation binders

A third type of ink is an evolution of thin film sol–gel processing—often termed composite sol–gel [96,108,109]. Unlike the fugitive binder inks, these transformation binders use a metallorganic system (e.g. sol–gel) to replace the organic binder and carrier fluid. In these systems, following deposition and evaporation of the solvent, the sol–gel material gels and binds the ceramic particles to each other and to the substrate. Subsequent heating then causes the sol to transform to an amorphous oxide ceramic and subsequently to a polycrystalline ceramic, which fully integrates with the ceramic powder. Due to the high reaction kinetics of the sol gel material it is possible to produce thick films at temperatures as low as 600 °C. The density of the films can also be varied by changing the proportion of sol to powder present in the ink and through using repeated infiltration steps [94] to enhance the green density.

As with the fugitive binder inks, sintering aids can also be added to further enhance to the densification of the systems [80,94,97]. As it is necessary to melt the sintering aids, typical processing temperatures are above 700 °C when sintering aids are used.

In addition to the ability to process films at lower temperatures it is also possible to create composite structures where the composition of the powder is different to that of the sol–gel material [80], which offers the potential to tailor the behaviour of the film.

A variation of the composite sol–gel route is to use conventional thick film processing followed by infiltration of the oxide ceramic producing sol to increase the green density of the film [110]. Subsequent sintering is then enhanced due to the presence of the highly sinterable sol–gel derived nanophase.

2.2.4. Electrical properties of PZT thick films

As the functional properties of ceramics depend strongly on the composition, dopant level, film stress state, film thickness, processing temperature and substrate materials it is very difficult to find meaningful comparisons from literature. In general, it can be stated that film properties are relatively independent of the route used to produce the films. The major factors affecting the properties are the sintering aid—with lower properties being exhibited as the percentage of sintering aid increases above the optimum level (typically 2–10 vol%). For a given sintering temperature, better properties are found in films containing the sintering aid with the lower melting point or activity.

When compared to comparable bulk ceramics, the relative permittivity of thick films can approach that of the bulk for certain compositions. However, piezoelectric and ferroelectric properties are typically reduced to between 1/4 and 1/3 those of the bulk ceramic [111,100,110,112] due to a combination of domain clamping and the presence of secondary phases.

2.2.5. Summary—ferroelectric thick films

A number of thick film deposition techniques have been presented which can themselves be used with different types of ink. Through the appropriate selection of process and ink it is possible to tailor the process for a wide range of applications.
Very low temperature (<100 °C) processing can be accomplished using paint type systems. Such systems exhibit very much lower electromechanical properties than do bulk ceramics. To improve the functional properties of the system it is necessary to use higher processing temperatures (650–950 °C). Such processing temperatures require the use of a liquid phase sintering aid or sol–gel material to lower the processing temperature from 1200 °C typically used for processing functional ceramics. The reduction in temperature means that these materials can be successfully integrated into microsystems.

3. Piezoelectric semiconductors

Piezoelectric semiconductors, including the Groups III–V materials such as GaAs (see Table 1), have attracted growing interest for MEMS applications where one particular advantage is their compatibility with conventional processing technologies for integrated circuit technology. Unlike perovskite materials such as PZT described earlier, the materials are not ferroelectric in nature and cannot be poled. Therefore, interest in these materials is primarily in the form of thin films that are epitaxially grown on a substrate. Although reported research [113,114] on specific semiconductor materials such as those with the noncentrosymmetric wurtzite crystal structure (Fig. 4), such as GaN, ZnO, CdS and ZnS, often refer to their ‘high’ piezoelectric coefficients it should be stressed that the coefficients are considerably lower than those of ferroelectric ceramics such as PZT. Nevertheless, published research on MEMS based devices using the piezoelectric or pyroelectric properties of such materials is growing and the aim of this section is to provide an overview of the most common materials and their potential advantages and applications. Table 1 presents a summary of dielectric, piezoelectric and mechanical properties of a range of piezoelectric semiconductors, along with the properties of a PZT for a direct comparison. The $d_{ij}$ coefficients (strain per unit field) are low compared to PZT indicating a low level of strain for actuator applications. It is of interest to note the much lower relative permittivity of these materials compared to PZT, which can lead to high $g_{ij}$ coefficients ($g_{ij} = d_{ij}/\varepsilon$), which is measure of the electric field per unit stress for sensor applications.

3.1. Groups III–V nitrides (GaN/AlN)

The piezoelectric effect in III–V nitrides, such as GaN, has been of primary interest because of its influence on the behaviour of field electric transistors and light emitting diodes [113]. However, interest has grown in utilising the piezoelectric coefficients of these materials to develop MEMS systems, sensors and actuators which can take advantage of the inherent wide band gap (3.4 eV for GaN), chemical and radiation inertness and high temperature properties of GaN. The wide band gap makes it a candidate material for high-power and high-temperature or radiation resistant electronics, particularly above 180 °C, which can degrade conventional silicon based transistors. Depending on whether the growth of GaN is in the [0001] direction (when its crystal structure is wurtzite) or in the [111] crystal direction for the zinc blend crystal structure, the material exhibits a strong lattice polarisation, which has been reported to be ideal for piezoelectric based transducers or pyroelectric sensors at temperatures in excess of 300 °C [113].

Shur et al. [113] recently reported data on the performance of preliminary devices based on GaN. The pyroelectric effect in GaN thin films was used to generate an electric charge in response to flow of heat at temperatures up to 300 °C. In addition to the primary pyroelectric effect the secondary pyroelectric effect was considered where a charge is developed as a result of a piezoelectric induced thermal strain. For fast heat transfer, the primary effect was dominant and experimentally determined pyroelectric coefficients for n-type GaN were $\sim 10^4$ V/mK, comparable to typical pyroelectric ceramics such as PZT and BaTiO$_3$. Modelling results, with electrode contact along the $c$-axis, predicted that the potential sensitivity of GaN could be increased up to $7 \times 10^5$ V/mK, which is greater for other high temperature ferroelectric pyroelectric materials such as LiTaO$_3$ [113].

Relatively little work has examined the development of strain sensors or actuators based on GaN. Strittmatter et al. [114] stated that one possible reason for this is that, compared to conventional ferroelectric ceramics, the movement of free charge carriers in GaN can potentially negate any piezoelectric charge that is developed as a result of a force or strain. To minimise this effect, Strittmatter used Schottky diodes on n-GaN to hinder the screening effect and examined the electrical response of a GaN cantilever as a result of an applied strain. It was concluded that the voltage generated

---

19 Christopher R. Bowen
Table 1
Collated data of piezoelectric materials

<table>
<thead>
<tr>
<th>Material</th>
<th>GaN</th>
<th>AlN</th>
<th>CdS</th>
<th>ZnS</th>
<th>GaAs</th>
<th>ZnO</th>
<th>InN</th>
<th>PZT5H</th>
</tr>
</thead>
<tbody>
<tr>
<td>e33 (C/m²)</td>
<td>1 [113]</td>
<td>1.55 [113]</td>
<td>1.32 [134]</td>
<td></td>
<td>−0.185 [113]</td>
<td>1.32 [113]</td>
<td></td>
<td>0.43 [113]</td>
</tr>
<tr>
<td>e31 (C/m²)</td>
<td>−0.36 [113]</td>
<td>−0.58 [113]</td>
<td>−0.074 [134]</td>
<td></td>
<td>0.093 [113]</td>
<td>−0.57 [113]</td>
<td>−0.22 [113]</td>
<td></td>
</tr>
<tr>
<td>e15 (C/m²)</td>
<td>−0.3 [113]</td>
<td>−0.48 [113]</td>
<td></td>
<td></td>
<td>0.093 [113]</td>
<td>−0.48 [113]</td>
<td>−0.22 [113]</td>
<td></td>
</tr>
</tbody>
</table>

In some cases data from the stiffness matrix [C] is shown rather than compliance [S].
was a direct result of the piezoelectric effect and was not piezoresistive in nature. The use of GaN solid state sensors for harsh environments, including space, aerospace and homeland defence has been reviewed by Pearton et al. [115]. Free standing GaN cantilevers have been produced on silicon substrates [116]. Little data is provided on GaN actuation, apart from extensional measurement of films to characterise the relatively small $d_{33}$ coefficients (compared to PZT) of these materials [117].

Another wide band gap (6 eV) group III–V nitride, AlN, has been considered for surface wave technology in thin film form due to its high piezoelectric coefficients for MEMS applications. Low deposition temperatures ($<500$ °C) enable compatibility with conventional integrated circuit technology [118,119] and the large band gap provides high resistivity that minimises dielectric losses [128]. Iborra et al. [119] examined the use of RF-sputtering to produce thin films with low residual stress and high c-axis orientation to optimise the piezoelectric properties. Preliminary test structures consisting of AlN/polysilicon bimorphs were fabricated, although no actuation results were reported. Cochran and co-workers [120] recently reported the use of thick AlN films (>5 $\mu$m) for bulk acoustic wave resonators, with its high Curie temperature, low permittivity and low losses being cited as potential advantages for such applications.

### 3.3. ZnO materials

Zinc oxide is another semiconducting and piezoelectric material that has potential applications in optoelectronics, piezoelectric sensors, resonators, surface acoustic wave devices and transducers [123,124]. Examination of Table 1 reveals that it has relatively high piezoelectric coefficients, along with its wide band gap (3.4 eV), near UV emission and transparency to visible light. A range of piezoelectric nanostructures including springs, helices, rings and bows can be formed due to polar-surface dominated growth mechanism of ZnO. Fig. 5 shows some of the ZnO nanostructures, synthesised by thermal evaporation of solid powders. Its relatively high piezoelectric coefficient (Table 1) provide a means to develop electro-mechanical coupled sensors and actuators at the nano-scale [123,124]. It has been reported that the piezoelectric coefficient of a polar nanobelt is approximately three times greater than that of the bulk material [124].

ZnO has been also used in more conventional thin film form (~0.5-μm thick), sputtered onto a 1.5 $\mu$m silicon nitride membrane, to develop a piezoelectric micromachined membrane acoustic devices, e.g. miniature microphones or loudspeakers for a cellular phone or earphone [125]. The maximum displacement of the membrane was 1 $\mu$m at 7.3 kHz with an input drive of 15 $V_{0-P}$. For this material, cantilever actuators based on ZnO have been manufactured and characterised. DeVoe and Pisano [126] produced 500 $\mu$m long ZnO cantilever actuators by surface
micromachining to validate a novel modelling methodology for cantilever actuators, with experimentally measured deflections of up to 400 nm. Minne et al. [127] produced cantilevers for atomic force microscopy, with a deflection of over 4 μm at dc conditions and 30 μm at resonance with applied electric fields up to $10^7$ V/m. Trolier-McKinstry and Muralt [128] have provided a good overview of ZnO thin film MEMS, along with a comparison and discussion of AlN and PZT based films, along with the appropriate figures of merit to compare materials for particular applications.

3.4. Summary—piezoelectric semi-conductors

Piezoelectric semiconductors such as those in Table 1 are attracting growing interest for MEMS applications due to their compatibility with conventional processing technologies for integrated circuit technology and wide gap for use in harsh environments. The piezoelectric ‘d’ coefficients are lower than many ferroelectric materials and research on actuator applications to date is limited. The low permittivity results in high ‘g’ coefficients indicating potential for sensor applications.

In addition to piezoelectric properties, materials such as ZnO nano-rods can be employed as a bio-chemical sensor to improve the physiological sensors sensitivity and selectivity. This topic is discussed in the following section.

4. Zinc oxide structures for chemical sensors

In addition to its useful piezoelectric properties, zinc oxide has attracted interest for applications in chemical sensing. There is an increased demand for selective, sensitive, time domain chemical sensors for physiological...
environments, primarily due to the interest in human health care and the need for new drug discovery. Almost all chemical and biochemical reactions involve a process where the acidity (pH) is subjected to relatively small changes, sometimes, even momentarily. In real physiological mediums, the problem is made complicated by the fact that the pH changes have to be detected in volumes that are relatively small. By implication therefore new sensor systems must also be small to be effective. In general, when objects are scaled down isomorphically (i.e. all dimensions are scaled uniformly) the change in length, area and volume ratios increase as we scale down and this renders surface effects to be significant. This alters the relative influence of the different physical effects in question in an unexpected way. If the object (e.g. analyte) in question shrinks to the same length scale as the effective boundary layer then continuum theories break down and the laws of micro scaling no longer apply. For the analyte in question, the total sample size needed for the detection is determined by the analyte concentration [138]. An important property of over-riding interest is the achievable sensitivity and sensors with a wide dynamic range of detection sensitivity are required in most cases. Before proceeding, we define the sensors of interest here to be those called electrochemical sensors. It is important to mention that electrochemical sensors are more flexible to miniaturization and usually provide a large dynamic range. They are further divided into conductometric, potentiometric and amperometric. An electrochemical sensor is a sensor that deals with the electron transfer, electron consumption or generation during a chemical or biochemical process. It is also important to note that, a potentiometric sensor measuring a voltage such as the ion sensitive filed effect transistor (ISFET) or ion selective electrode (ISE), are scale invariant; while amperometric and conductometric sensors on the other hand measures currents and are sensitive to miniaturization. The reduction in sensor size can lead to beneficial effects. To illustrate this, we consider the sensitivity of a sensor as we miniaturize our electrodes. If the size of the sensing electrodes is reduced to sizes comparable to the thickness of the diffusion layer, and the electrodes are kept isolated, non-linear diffusion, caused by curvature effects, needs to be considered. Analysis of such a situation showed that as the non linear curvature effects become more and more pronounced, more diffusion takes place, i.e. diffusion occurs from all directions and analyte collection increasingly persist. This leads to more analyte supply to the electrode; this is an example of a beneficial un-expected effect. This implies that as we scale down our sensing electrodes and keep them “isolated” the sensitivity is significantly enhanced. In fact, it has been demonstrated experimentally that if a single ion is located near a single electron transistor (SET), detection is achieved with a change of the conduction current. In addition to this, nano-electrodes have relatively large surface area that makes them attractive for pH and other chemical sensing. In addition, the possibility to control their nucleation sites makes them one of the best candidates to develop high-density sensor arrays. Miniaturization is also a mixed issue for both the sensor and analyte. Hence, the analyte concentration is an important parameter to consider. As mentioned above, sometimes we are faced by the fact that the analyte to be detected has relatively low concentration (≈fL), and this implies the need for a large sample volume to achieve detection. Large volumes are again not in our control, especially if we deal with a real physiological medium, e.g. human body analyte.

In this section, we will briefly discuss the properties and use ZnO nano-rods (with few nanometers in diameter and micrometers of length) for chemical sensing purposes [139]. Experimental results from growth as well as theoretical results on sensing using different approaches will be presented.

4.1. Synthesis and properties of ZnO nano-structures

Zinc oxide (ZnO) is a direct band gap semiconductor (3.37 eV at room temperature) and having large exciton binding energy (60 meV), exhibiting near UV emission, transparent conductivity and piezoelectricity [140–143]. Of interest to this section are the bio-safe and bio-compatible properties of ZnO. In addition, ZnO is a polar semiconductor, which means that the outer surface can be controlled to have a neutral, positive (Zn$^+$ terminated), or negative charge (O$^-$). This, combined with the bio-safe and biocompatible properties, indicated that ZnO is a suitable material for chemical sensors in physiological media.

A variety of ZnO nano-structures (nanometer of diameter and micrometer of length) have been synthesized using different techniques. Nano-structure geometries include nano-rods, nano-wires, nano-belts, nano-rings and nano-tubes [144]. The most commonly investigated growth technique is through vapor phase nucleation. The vapour species are first generated by evaporation, chemical reduction and gaseous reaction. They are then transported and condensed onto the substrate. This is illustrated in Fig. 6. Fig. 7 shows an example of well aligned ZnO nano-wires grown on sapphire [145]. Although this technique has been extensively studied the exact mechanism of growth by the vapour
phase technique is, as yet, not well understood [146]. An alternative technique for growing well-controlled ZnO nanostuctures is by hydrothermal reaction [147–150].

The unique characteristics of nanoscale materials make them a perfect choice for the sensors world. Integrating them into existing sensors can increase the device sensitivity, selectivity, and speed. In addition, the large surface area to volume ratios greatly facilitates sensor miniaturization with benefits discussed above. Because of the minuscule size of nanoscale materials, their chemical and physical properties differ from those of their bulk counterparts and they therefore behave differently. One obvious benefit is their ability to be functionalized or custom-designed to attract specific molecules; another is their extremely high surface area, contained within a tiny effective volume. Researchers are integrating functionalized ZnO nanorods for a variety of sensor applications to meet urgent needs in fields ranging from biomedicine to biochemistry [145,151–154]. The goal is to lay the foundations for a miniaturized sensor that uses the smallest possible sample size to detect the smallest concentration of molecules of interest.

Since ZnO nanorods are conductive, they can provide a signal each time a target substance (H⁺ proton) binds to the surface layer of the nano-rod [145–147]. This provides advantages for both of the principle detection techniques that are used, namely the electrochemical potential method and the site-binding method.

### 4.2. Electrochemical potential method

When a solid is immersed in a polar solvent or an electrolyte solution, a surface charge is developed through one or more of the following mechanisms:

1. Preferential adsorption of ions.
2. Dissociation of surface charged species.

![Fig. 6. Schematic diagram of a typical chamber of the synthetic growth of ZnO nanostructures.](image)

![Fig. 7. SEM images of c-axis oriented ZnO nano-wires grown on patterned sapphire substrate: (a) a low magnification image showing nanowires grown in squares and (b) higher magnification image showing nanowires within one square [145].](image)
(3) Isomorphic substitution of ions.
(4) Accumulation or depletion of electrons at the surface.
(5) Physical adsorption of charged species onto the surface.

For a given solid surface in a given liquid medium, a fixed surface electrical charge density or electrode potential, $E$, will be established, which is given by the Nernst equation:

$$E = E^0 - \frac{RT}{nF} \ln \left( \frac{f_{\text{production}}}{f_{\text{reaction}}} \right) + \frac{2.303RT}{nF} \log(a_{H^+}) - E_{\text{ref}}$$

where $E^0$ is the standard electrode potential of the zinc oxide redox probe, $F$ the faraday constant (96.5 kC/mol), $T$ the absolute temperature (298 K), $R$ the gas constant (8.314 J/mol K), $n$ the number of electrons in the redox reaction, $a_{H^+}$ the concentration of protons, $[\text{production}]$ and $[\text{reaction}]$ are the concentrations of the species, and $f_{\text{production}}$ and $f_{\text{reaction}}$ are their related activity coefficients.

The surface charge in zinc oxide is mainly derived from preferential dissolution or deposition of ions. Ions adsorbed on the solid surface determine the surface charge, and thus are referred to as charge determining ions. In the oxide systems, typical charge determining ions are protons and hydroxyl groups and their concentrations are described by pH. As the concentration of charge determining ions varies, the surface charge density changes from positive to negative or vice versa. The concentration of charge determining ions corresponding to neutral or zero-charged surface is defined as a point zero charge or zero-point charge (p.z.c.). At pH $>$ p.z.c., the oxide surface is negatively charged, since the surface is covered with hydroxyl groups, $\text{OH}^-$, which is the electrical determining ion. At pH $<$ p.z.c., $H^+$ is the charge determining ions and the surface is positively charged.

Zinc oxide is an amphoteric oxide which reacts with both strong acids and bases and it can thereby exhibit both acidic and basic properties (such materials do not usually react with water at all). In general to operate in an aqueous environment, the metal atom in an amphoteric oxide must be sufficiently electropositive to provide the oxygen with enough negative charge to strip a proton from a neighboring $\text{H}_3\text{O}^+$. However, the metal ion must also be electronegative enough to serve as an electron acceptor from a neighboring $\text{OH}^-$ (see Figs. 8 and 9) [155].

### 4.3. Site binding method

Measurement of the concentration of $H^+$ ions can be performed by the site-binding method [156]. It is assumed that the surface provides discrete sites for chemical reactions when it is brought in contact with an electrolyte solution. This method is very suitable for pure inorganic oxides ($\text{ZnO}$) where only a single type of site is present that has an amphoteric character. The sites can protonate or deprotonate leading to a surface charge and surface potential which is

![Fig. 8. Schematic diagram shows the intracellular electrochemical potentiometric sensor for physiological medium.](image)
dependent on the pH of the electrolyte (this means that each surface site can be neutral, act as a proton donor (acidic reactions) or act as a proton acceptor (basic reactions)).

The surface potential between the sensitive layer and electrolyte interface based on the site binding model is:

$$2.303(p_{\text{pzc}} - \text{pH}) = \frac{q\psi}{kT} + \sinh^{-1}\left(\frac{q\psi}{kT\beta}\right)$$

where $k$ is the Boltzmann’s constant, $k = 1.38 \times 10^{-23}$ J/K, $q$ the electronic charge $q = 1.602 \times 10^{-19}$ C, $\beta$ is the parameter, which can be expressed in terms of the acidic and basic equilibrium constants of the related surface reactions.

$$\beta = \frac{2q^2N_S(K_aK_b)^{1/2}}{kTC_{\text{DL}}}$$

where $K_a$ and $K_b$ are the dissociation constants. $N_S$ is total number of sites per unit area (density of surface OH groups). $C_{\text{DL}}$ is the double layer capacitance at the interface given by the Gouy-Chapman-Stern model [157]. The measured pH response, based on the site binding model, is dependent on the nanorod size. Thus, different nanorod sizes will give different of parameters $p_{\text{pzc}}$, $K_a$, $K_b$ and $N_S$ (see Figs. 10 and 11).

Fig. 9. Relationship between the final potential of intracellular sensor (in volt) with change of pH of two different nanorods ZnO ($r = 200$ nm, $l = 10$ μm) and ($r = 50$nm, $l = 2$ μm). The upper two curves represent the basic reaction and the lower two curves the acidic reaction.

Fig. 10. Schematic diagram of electrical ionic double layer close to the ZnO nanorods surface.
ZnO nano-rods can be employed as bio-chemical physiological sensors with improved sensitivity and selectivity. They can be chemically customized to suit a wide variety of applications. With their ability to react rapidly and with extreme sensitivity such new materials may dramatically improve sensing technology and in combination with other functional multi-scale materials, they can open many new opportunities for sensors in biomedical applications.

5. Silicon carbide for chemical sensing devices

Silicon carbide (SiC) is the only compound that exists in the Si–C two atom system. However, it exists in more than 180 poly-types. These all consist of identical closely packed Si–C double layers, whose stacking sequence differs along a certain direction. The nearest neighbor arrangement of atoms is identical in the crystal structures. Each carbon atom is tetrahedrally surrounded by four Si atoms, and each Si atom is tetrahedrally bonded to four carbon atoms by sp$^3$ hybrid orbitals [158]. The ionicity of SiC is 12%. The next nearest neighbors may be placed in two different possible ways. These are respectively the cubo-octahedral or the hexagonal cubo-octahedral. Single crystalline SiC exists in cubic (C), hexagonal (H) and rhombohedral (R) structures. Moreover, all the poly-types are divided into two families. The $\alpha$-SiC and the $\beta$-SiC families. The $\beta$-SiC family has only one member and it is the only poly-type that exists in the cubic structure, referred to as 3C-SiC (three layers cubic). It is the simplest and the most well known among all the poly-types. Despite the large lattice mismatch with Si (20%), the cubic structure is in fact important since it enables the materials to be grown epitaxially on Si. The growth of a device quality 3C-SiC/Si heterostructure presents the possibility of monolithic integration with other standard Si electronic components. The other poly-types, belonging to the $\alpha$-SiC family exist in the other two crystal structures mentioned above. Amongst these, the 2H, 4H, 6H and 16R are the most commonly occurring poly-types. The number denotes the number of the repeated layers, which in some poly-types reaches a few hundreds of repeated sequences. Although the chemical bonding and thermodynamics of the different poly-types seem to be the same, some of the physical properties such as bandgap and electronic properties differ very strongly. This implies that we can consider all of the poly-types as belonging to the same family of semiconductors having almost the same lattice parameters and similar chemical properties, but with different physical and electronic properties. This is of great interest as the combination of different SiC poly-types can lead to heterojunctions among the same family. In general all the poly-types of SiC are characterized physically by a wide bandgap. The 2H poly-type has the largest bandgap (3.33 eV), while the 3C poly-type has a bandgap of (2.39 eV). Beside the large bandgap, SiC is an excellent radiation-resistant material, having a high Debye temperature and high thermal conductivity. It is important to also note that it is straightforward to grow high quality oxide on SiC. These properties make SiC an attractive semiconductor material for many application areas, which include, power electronics, high frequency devices, and sensors. The later is of interest to this section. Comparing the physical properties of the different wide bandgap semiconductors, it can be concluded that there is no competition based on

---

21 M. Willander, O. Nur, Q. Wahab
physical properties alone. The physical properties of all wide bandgap semiconductors are similar, apart from the offered direct and indirect bandgap which influences the quantum efficiency in optoelectronic devices. Significantly, SiC demonstrates high chemical stability and it is a chemically an inactive material, making it a durable and appropriate choice for sensors operating under harsh conditions, e.g. operating at high temperature. High chemical stability also suggests that this material can be used for applications in medicine and ecological engineering fields and SiC has long been recognized as an excellent material for the realization of Micro-Electro-Mechanical Systems (MEMS) operating in harsh conditions. This is due to its unique mechanical strength combination with chemical inertness [159]. In this section, we will restrict the discussion in the sections below towards chemical sensors.

The growth of single crystalline SiC and device processing technology research has been going on for few decades and remains an active area of research. It has attracted many research laboratories around the world since the mid-80s. The growth of different poly-types (both bulk as well as thin films) has reached a device quality material for many applications. Since semiconductor crystal quality is an important factor we will discuss some important facts regarding SiC single crystal growth. We first briefly describe the growth development of single crystalline SiC (bulk and thin films), then the general principles of chemical gas sensors are mentioned. This will be followed by some specific examples that utilize SiC as a chemical sensor. Finally, some ‘exotic’ chemical sensors based on new and/or innovatively engineered SiC will be given.

5.1. SiC single crystal growth

As mentioned above, the possible device applications utilizing SiC are many and, depending on the application, there is an acceptable degree of crystalline quality for efficient and long life time component. Fig. 12 shows a high resolution cross-sectional transmission electron micrograph (HR-TEM) from a 4H-SiC along the 0001 plane. The image shows the presence of stacking faults that can lead to current degradation with time. The presence of such stacking faults is the major reason for the lack of availability of commercial SiC devices for high power applications.

The sublimation technique for single crystalline semiconductors is one of the oldest growth techniques used for bulk SiC [160]. These initial efforts resulted in high purity α-SiC single crystals. However, the crystals were random in size and the control over the grown poly-type was poor. This technique was later modified to include seeded growth [161], which resulted in more control of the particular poly-type grown and was adopted by many researchers. Using the seeded sublimation technique, large single-crystal boules of pure 6H-SiC poly-type with diameter up to 40 mm as well as 4H-SiC were demonstrated [162,163] and doping of this large single crystal to tailor electrical properties was possible. Although the modified seeded sublimation growth method has demonstrated control over the poly-type, along with possibility of obtaining lightly doped and highly doped crystals, the structural quality regarding the density of dislocations was not promising. A severe problem was the density of micropores and micropipes (around $10^2$ cm$^{-2}$). This density of defects is a severe limitation on the use of these crystals and it can cause degradation in

![Fig. 12. High resolution cross sectional transmission electron microscope (HR-TEM) view of 4H SiC along the [0001] plane.](image)
device performance, especially for high power electronics. Many other techniques were proposed to grow well controlled high quality single crystal SiC and chemical vapor deposition (CVD) has been recently favoured. It is the most advanced epitaxial technique and is the most widely used technique for commercialization [164]. As mentioned above the 3C-SiC is of special interest due to the possibility of epitaxial growth on Si and hence monolithic integration with other Si standard devices. Although the lattice mismatch between 3C-SiC and Si is relatively large, different methods have been employed to produce device quality 3C-SiC/Si heterostructures. Some early results can be found in [164–166]. One approach adopted to reduce the lattice mismatch, is to use off-axis substrates to eliminate the anti-phase boundaries which are the main type of intrinsic defect in the grown 3C-SiC layers [167]. A second approach is the use of elastic substrates to release stress due to the relatively large lattice mismatch [168]. It should be noted that elastic substrates, e.g. silicon on insulator (SOI) becomes viscous at temperatures approaching the growth temperature and hence their use will not lead to any temperature modification.

The first commercial SiC substrates became available during the late 80s and early 90s. This has stimulated the research on SiC-based device technology and triggered a global interest. The status of SiC single crystal grown epitaxial layer today is to a large extent acceptable for almost all applications. Until 5 years ago the density of dislocations, micropipes and micropores, although low, were still unacceptable. Today these un-wanted structural imperfections are completely eliminated and only few stacking faults (as those shown in Fig. 1) are present. It is important to note that although stacking faults can cause degradation in device performance for high power applications, they are not of significant concern when considering SiC chemical sensors.

Although SiC based devices demonstrate significant advantages and potential the projected realization for commercial products has not matched initial expectations. This is due to many factors. For the specific case of SiC based chemical sensors, although stable performance up high temperatures has been demonstrated, mounting and packaging of the sensor is providing a further challenge for a wide range of commercial products. This issue is elaborated below when discussing specific SiC chemical sensors.

5.2. Gas sensor principles

The first pioneering work on chemical sensors was demonstrated during the early 70s. Two main different devices were demonstrated, these were the ion sensitive field effect transistors (ISFET) for pH sensing of electrolytes [169], and Pd gate MOSFETs for gas sensing [170]. Of particular interest here is the gas sensor MOSFET based on pure Si Pd-gate MOSFET. The principle of operation was based on the fact that Pd is a catalytic metal that dissociates the ambient gas to ions. These travel by diffusion to the metal-oxide interface where an electrically polarized layer is formed (for the first Pd MOSFET hydrogen was the ambient gas). This layer stimulates a change in the electrical characteristics of the MOS device, and hence a sensing mechanism is established. In the case of MOS based sensors, this is observed as a shift of the C–V characteristics due to the voltage modification of the dipole layer. Demonstration of this first solid state electronic chemical sensor has stimulated the scientific community and many research papers are published. Gas sensitivity and operating conditions using different catalytic metals were also demonstrated as will be discussed in the next section. Based on the same principle, two other devices employing SiC were demonstrated with excellent performance. These are the metal insulator semiconductor schottky diode (MIS) and the Schottky diode. The MIS Schottky usually employs a thin (0.15–0.2 nm) silicon dioxide layer to avoid pinning the barrier height.

5.3. SiC gas sensor development

Although the successful demonstration of the first gas sensor using pure Si found many application areas; Si-based MOS sensors cannot operate at temperatures above 280 °C.

Many gaseous systems of interest have an operating environment well in excess of the Si working temperature. High temperature gas sensors, for example, are of great interest for large scale commercial applications such as continuous monitoring of exhaust gases emerging from combustion processes in car engines. The requirements here are very demanding. To monitoring car exhaust gases, an efficient sensor will have a working temperature as high as 900 °C for 4000 h or alternatively 160,000 km. The response to a change between oxidizing and reducing atmospheres must be within 10 ms. SiC with its large bandgap and chemical inertness is the best among the wide bandgap semiconductors for this type of application.
The first MOS SiC based sensor was demonstrated on commercial 6H-SiC with Pt used as the catalytic gate metal [171,172]. This sensor has been operated at temperatures as high as 800 °C. The sensitivity to hydrogen was studied over a wide temperature range. Although MOSFET based devices have been demonstrated employing both the 4-H and 3C families, 3C-SiC has a higher electron mobility. Hence, it has the potential to provide a more efficient MOS device. However, the lack of 3C-SiC wafers (substrates) has led to high temperature sensor research focusing on the H family. The earliest reported SiC MOS-based sensors showed promising results and they were used to study the reactions of catalytic metals at temperatures well above the operating temperature for Si MOS (280 °C). The sensors enabled the study of the hydrocarbon dissociation, which usually occurs at temperatures above 350 °C. The SiC based MOS sensor has provided detailed information on the decomposition of different hydrocarbons, e.g. methane, ethane, propane and butane [172]. Since this early work, different groups have presented investigations that provided knowledge of high temperature chemical gas sensing and recent studies can be found in refs. [173–177]. The development of SiC based devices has shown promising results with stable operation demonstrated for temperatures up to 600 °C. The effects of operating temperature, catalytic metal employed and physical structure of the sensor on selectivity and sensitivity have been investigated for many gaseous environments. As a result of these valuable efforts, sensor arrays can be fabricated for the purpose of sensing gas mixtures. However, it should be noted that at high temperatures the band offset between the SiC and the SiO2 decreases, and as a result the devices are increasingly susceptible to charge injection from the semiconductor into the oxide. This is understood to be the main reason for sensor failure at high temperatures (above 600 °C). By engineering the physical parameters of the device in such a way to minimize the electric field at the SiC–SiO2 interface, stability above 600 °C can be achieved. This issue is important with regard to the long term stability of these sensors. The use of nano-particles embedded in the oxide has been suggested as a possible method for increasing the long term stability at relatively high temperatures. In addition to the SiC MOS sensor, MIS Schottky and Schottky fabricated on SiC were employed as gas sensors. Early work on these devices can be found in refs. [178,179]. It is important to note that Schottky contact is a majority carrier mechanism and hence a fast response is expected. The Schottky based sensor is also the most simple structure to fabricate. Fig. 13 shows a typical Schottky based SiC based high temperature gas sensor. Both devices show stable operation, but at temperatures below the maximum theoretical temperature associated with SiC physical properties. More device engineering research is needed in order to have stable operation at extremely high temperature.

5.4. Other innovative SiC based chemical gas sensors

We have provided a brief example of the use of SiC in an innovative way to demonstrate a highly sensitive sensor. It is known that as the sensing element dimensions are reduced, the more sensitive is the device. Most recently, an innovative sensor based on porous SiC was demonstrated [180]. This was sensitive to ammonia for levels as low as 0.5 ppm. The sensor was fabricated employing a relatively very simple process. SiC thin film was deposited on Si

![Fig. 13. Schematic diagram showing a typical high temperature Schottky based gas sensor. As seen the processing steps required to such sensor are quite simple. Here the top contact choice can be any appropriate catalytic metal depending on the sensitivity.](https://example.com/fig13.jpg)
using chemical vapor deposition. The SiC was made porous by electrochemical etching. Finally, Al electrodes were deposited to contact to both the porous-SiC and back side of the Si wafer. The exact sensing mechanism of this device is not yet completely understood, however one possible explanation could be due to changes in the depletion layer upon decomposition of ammonia and the subsequent adsorption of the hydrogen.

5.5. Conclusions

On carbide has been demonstrated to be the best candidate for high temperature chemical gas sensors. Progress in crystal growth and processing methods over previous years has ensured the availability of device quality wafers and thin epitaxial layers with precise control of doping, oxidation and metallization. The wide bandgap, combined with chemical inertness results in SiC being the best material for gas sensing in harsh environments or at high temperatures. The physical properties of SiC enable it to function at temperatures as high as 1000 °C. However, testing in research laboratories has shown that long term stability is currently restricted to ~600 °C. Engineering improvements to the sensor physical design would certainly push the operational performance up to much higher temperatures. Further development work is also required in the areas of mounting and packaging of this type of sensor system.

6. Magnetostrictive thin films

Interest in magnetostrictive thin films has rapidly grown over the last 10 years due to their potential as actuators for powerful transducer systems in microsystems [181] or as miniaturized sensors monitoring strains [182] or magnetic fields using multiferroic composites [183]. These developments are based on the direct magnetostrictive (Joule) or indirect magnetostrictive (Villari) effect, respectively. Fig. 14 shows the principle of both effects in the case of thin films with an in-plane anisotropy perpendicular to the applied field or stress.

While the direct effect results in a change of dimension of the material due to domain alignment, which is the origin of any solid state actuator, the situation for the inverse effect is more complicated since it may consist of 180° domain walls, as shown in Fig. 14. In that case, it is obvious that the strain or stress does not result in a net magnetization of the sample but may lead to a rotation of the magnetic domains. To use this effect as a sensor mechanism for mechanical quantities it is necessary to combine the inverse magnetostriction with further effects being themselves sensitive to the orientation of magnetic domains. The effects to be combined with inverse magnetostriction are magnetoresistive effects, especially GMR (giant) or TMR (tunnel) effects [184], magnetoimpedance [185] or inductance using microinductors [186,187].

In the following sub-chapters, the materials development for giant magnetostrictive materials, their use as micro-actuators, and magnetostrictive sensor using inductive, impedance or resistive effects are described in more detail.

6.1. Giant magnetostrictive thin films

The development of giant magnetostrictive thin film materials at room temperature is based on rare earth-transition metal alloys. These alloys offer the best possibility to develop giant magnetostriction at room temperature or above, since the highly aspherical 4f orbitals of the rare earths, which are the origin of the large magnetostriction, remain in an oriented state due to the strong coupling between the rare earth and the Fe or Co moments. An important development task for giant magnetostrictive materials has been their optimization in terms of their magnetostriction to magnetic anisotropy ratio in order to attain large strains at moderate magnetic fields. In bulk materials, this was achieved by using cubic compounds, the rare earth-Fe2 Laves phases, in which the second order anisotropy constant vanishes along with Tb–Dy alloying to compensate the fourth order anisotropy constant [188]. In the case of thin films, amorphous (Tb,Dy)4(Fe,Co)1–x films [189] or particularly novel TbFe/FeCo multilayers [190] represent the most promising approaches to combine soft magnetic and giant magnetostrictive properties.

Since the magnetic saturation field is proportional to the ratio of the anisotropy and the saturation magnetization, two approaches are possible: the decrease of the anisotropy, e.g. by using amorphous magnetic materials or the
increase of the saturation magnetization. Using multilayers, it is possible to engineer novel composite materials which show enhanced magnetizations in comparison to classic thin film magnetostrictors [191]. To create such a material two materials have to be combined: one material is the giant magnetostrictive amorphous TbFe alloy which should be combined with a material that is magnetically soft with a very high magnetization and preferably a magnetostrictive (e.g. FeCo). By fabricating layers with thicknesses smaller than the ferromagnetic exchange length and domain wall width, domain wall formation at the interfaces is prevented. The magnetic properties of such an exchange-coupled multilayer system are determined by the average of each individual layer. Together with the reduction in anisotropy this increase in magnetization results in a significant reduction in the magnetic saturation field. Fig. 15 compares the magnetoelastic coupling coefficient \( b \) (defined as the product of magnetostriction and the effective Young’s modulus) of a TbFe/FeCo multilayer film with a nanocrystalline TbDyFe film and a TbFe single layer film.

The fabrication of giant magnetostrictive thin films has only been realized by PVD-techniques until now, the most prominently used being magnetron sputtering using either mosaic type or composite targets as well as multitarget arrangements. Other PVD-techniques include electron beam evaporation, laser ablation and ion beam sputtering.
The amorphous films are generally deposited onto unheated or cooled substrates while for crystalline films heated substrates for a single-step-process are used as an alternative to a post-deposition crystallization treatment.

For developing devices, the orientation of the magnetic easy axis and of the domains in the demagnetized state is of special importance since maximum magnetostriction is only obtained by 90° rotations of the magnetic domains, while 180° rotations do not result in any magnetostrictive strain. Considering that magnetostrictive thin film actuators are in general driven by a single magnetic field, whose direction is fixed in relation to the actuator and in-plane in order to avoid large demagnetization losses, the optimized demagnetized state should consist of domains with an in-plane easy axis being oriented perpendicular to the driving field direction (see Fig. 14). This initial state can be obtained by a post-deposition annealing process under an in-plane magnetic field which is oriented under 90° towards the driving field direction [192].

6.2. Magnetostrictive thin film actuators

Magnetostrictive thin films have been used as micro-actuators for applications in microsystem technology [181] predominantly based on a bending transducer principle consisting of a film/substrate bilayer with the substrate being in general non-magnetostrictive. The most important configurations have been cantilevers, membranes and plates, where magnetostriction in the film causes the film/substrate composite to bend, similar to the bending of a bimetallic transducer. Commonly, the most important feature of these micro-actuators is the possibility of a remote-controlled operation.

In the case of cantilever actuators, different applications have been realized, normally by using Si micro-machining for the fabrication of the cantilevers. The large bending or deflection of these cantilevers were used in the application of fluid jet deflectors controlling up to 500 ml/s [193], for magnetic field measurements by detecting the deflection of the cantilever optically, as well as for optical 2D micro-mirrors which employ bending and torsional vibrations driven by two differently oriented magnetic fields [194]. Furthermore, special attention has been paid to the development of thermal drift-free actuators either realized by a special design of the micro-machined substrate or by combining positive and negative magnetostrictive materials in a bimorph structure. Magnetostrictive membranes have been used for fluidic micro-components such as micro-valves or micro-pumps, whereas the deflection of the membrane is either used to close or to open the valve outlet or to induce a pressure rise in the pumping chamber [195]. Free plates with magnetostrictive coatings have been employed for the realization of traveling machines or for ultrasonic motors [196]. Both linear and rotating standing wave motors were realized with micro-machined Si or Ti substrates using a propulsion mechanism of vibrating teeth on a friction layer. The ultrasonic motors were operated by ac magnetic fields in combination with a magnetic bias field.

6.3. Magnetostrictive magnetoresistive sensors

The motivation to use the inverse magnetostriction for mechanical sensing is driven by their high sensitivity, small sizes, high spatial resolution, along with their cost-effective fabrication by thin film technology. These general advantages are especially fulfilled in the case of sensors combining magnetostrictive and magnetoresistive effects.

Giant magnetoresistive (GMR) or tunnel magnetoresistive (TMR) stacks are well known as highly sensitive magnetic field sensors [197,198]. The GMR or TMR structures consist in principal of a magnetic reference layer and a magnetic sensing layer which are separated by a metallic non-magnetic layer (typically Cu) in the case of GMR sensors and a very thin non-conducting layer (Al₂O₃ or MgO barrier) in the case of TMR sensors. For TMR sensors resistivity changes \( \Delta R / R \) of up to 350% have been obtained in the case of epitaxial MgO tunnel barriers between the high resistivity state (antiparallel orientation of the two magnetic layers) and the low resistivity state (parallel orientation). The GMR effect has been found in exchange-coupled structures such as Fe/Cr-multilayers with \( \Delta R / R \) as large as 220 at 1.5 K and 42 at room temperature [199].

Initial investigations that combine GMR structures with magnetostrictive phenomena have been limited to NiFe sensing layers of varying compositions [200–202]. A significant enhancement of stress sensitivity of GMR structures could be achieved by using highly magnetostrictive materials such as Fe₅₀Co₅₀ as a free layer [203]. The most important increase in sensitivity compared to magnetostrictive GMR sensors could be achieved by using TMR sensors in combination with highly magnetostrictive materials such as crystalline Fe₅₀Co₅₀ [204] or amorphous (Fe₉₀Co₁₀)₇₈Si₁₂B₆ [205].
For strain sensor applications the gauge factor (GF) which is defined as:

$$GF = \frac{\Delta R/R}{\Delta \varepsilon}$$

(3)

where \( R \) is electrical resistance and \( \Delta \varepsilon \) is the change in strain is the most important parameter for the classification of different sensors and demonstrates this enhancement. While FeCo/Cu/Co GMR-trilayer systems show gauge factors of approximately 30 (\( \Delta R/R \) of 3% upon 0.1% strain), the TMR sensor with Fe\(_{50}\)Co\(_{50}\) reaches values up to 600 (\( \Delta R/R \) of 25% upon 0.04% strain).

The potential application areas are manifold. The positive features are the very small size, for example TMR stacks have been demonstrated with junctions sizes as small as 50 nm \( \times \) 50 nm [206], the high sensitivity which also enables a reduction of the total sensor size and the large variety of substrates (e.g. Si, glass, diamond, polymer). The fabrication processes are in general compatible to both Si-micromachining and CMOS-technology offering the possibility to design highly integrated sensors.

6.4. Magnetostrictive magnetoimpedance sensors

These sensors profit from their high sensitivity, their possibility of remote interrogation [207,208], and the cost effective fabrication of melt-spun wires and ribbons. The giant magnetoimpedance (GMI) effect is widely considered as a high-frequency analogy of the GMR effect: a type of side effect of the skin effect. The sensitivity as a field sensor can reach 100%/Oe [209], which is typically at least one order of magnitude higher than that of GMR sensors. Sensors based on the GMI effect use magnetic ribbons or wires produced by melt spinning techniques [210], glass covered magnetic wires [211,212] or magnetic thin films [213–215]. In the case of magnetostrictive GMI sensors for stress/strain measurements, the magnetic material has either a positive or negative magnetostriction. As the skin depth is dependent on the permeability a change in permeability due to magnetostriction results in corresponding magnetoimpedance changes. Therefore, a general condition is that the skin depth is smaller than the cross sectional dimension of the material (wire, ribbon, film) in order to have a strong impedance dependency on changes in the skin depth.

GMI thin film micro-strain sensors were achieved using FeCoBSi materials. The sensor consists of a millimeter-sized meander-like structure on a 20 mm \( \times \) 4 mm and 150-\( \mu \)m thick glass cantilever. A change in impedance of about 46% was observed with 100 MHz driving frequency at a strain of 0.03% [214,216].

6.5. Magnetostrictive inductive sensors

Since one of the consequences of the inverse magnetostrictive effect is the change in permeability, inductive elements with magnetostrictive cores offer the possibility to produce magnetostrictive strain sensors. The application of a strain results in a change of permeability of the magnetostrictive core. This leads to a change of inductance \( L \) that, for example, can be monitored by electrical resonance measurements [187]. These measurements use resonant circuits (so called \( LC \)-tags, consisting of a capacitance \( C \) and an inductance \( L \)), where \( L \) is changed due to its magnetostrictive core. A particular benefit of this approach is the possibility to operate these sensors by remote interrogation using either radar reflectivity or inductive coupling.

Micro-inductors with incorporated magnetic thin films have been widely investigated in the past 10 years due to their potential as integrated components in micro-electronic devices. Four general types of magnetic thin film inductors [217] have been investigated in detail: strip line inductors with magnetic sandwich layers [218–220], spiral inductors with magnetic sandwich layers [221–223], solenoid inductors [224,225] and toroid inductors [226]. In principle, the same approaches can be also used in the case of mechanical sensors; the main difference is merely the replacement of the magnetic core by a magnetostrictive one. Another design aspect is the orientation of the magnetic easy axis which should be well-defined with respect to the direction of the applied stress in order to guarantee a reproducible and optimized sensor behavior. Therefore approaches which can be operated with a uniaxial anisotropy, as with a strip-line or the solenoid inductor are of more general interest for sensors.

An example of an inductive strain sensor was realized using a magnetostrictive \( LC \)-circuit based on a strip-line type inductor [227]. Here, the magnetic easy axis is aligned parallel to the strip, optimally aligned to the exciting high frequency magnetic field. When a tensile stress is applied perpendicular to the strip (for a positive magnetostrictive
material), the permeability and therefore the inductance of the strip will increase, resulting in a decrease of the resonance frequency. Similarly, a compressive stress leads to a stabilization of the easy axis and decreases of the permeability along the hard axis resulting in an increase in resonance frequency. Millimeter-sized LC resonant circuits have been fabricated in thin film technique on glass wafers [228]. FeCoBSi and FeCo/CoB were used as magnetostrictive materials, demonstrating both high frequency properties and high magnetostriction. The values of the inductance and the capacitance were designed to yield a LC resonance frequency of approximately 500 MHz.

It was demonstrated that this sensor can read by remote interrogation [229]. In this case, the read out was realized by two pick up coils and a network analyzer. The figure of merit (FoM) which was defined as

$$\text{FoM} = \frac{\Delta f / f_0}{\Delta \epsilon}$$

With $f_0$, $\Delta f$, $\Delta \epsilon$ being the resonance frequency and change in frequency and permeability with the change of strain, respectively. In the sensitive strain range this sensor reached a FoM of approximately 1000. These values are extraordinarily high compared to more traditional strain gauges. A possible application for this type of sensor is the measurement of torque. Torque measurements of up to 200 Nm using a similar sensor are presented in ref. [229].

7. Magnetic properties of magnetic nanoparticles

7.1. Single domains

This section will discuss the basic magnetic properties of magnetic nanoparticles and their various applications. Spherical magnetic nanoparticles with diameters smaller than about 30 nm, generally speaking, are single domains [230], although the critical single domain size depends on the material and the exact geometry of the particles. By definition in a single domain particle, the electrons that give the material its magnetic character are all spinning in the same direction, with the total magnetic moment being the sum of all the spins. Thus, the magnetic moment is proportional to the product of the particle volume and a material parameter called the intrinsic spontaneous magnetization of the single domain. Above the critical single domain size, the particles assume the multiple or ‘polydomain’ structure of bulk materials. Table 2 shows the critical single domain size and the intrinsic spontaneous magnetization for some typical magnetic materials.

The intrinsic spontaneous magnetization of single domains with sizes in the nanometre range is in most cases quite different to the corresponding value for the bulk material. This has been explained by a surface effect due to the effect of the small size of the nano-particles [231]. Surface effects become very significant in this case since the number of surface spins becomes comparable to the number of bulk spins for particles with sizes in the nanometre range. Spins at or close to the particle surface may become canted or tilted, so that they are not oriented in the same direction as the bulk spins. This results in an intrinsic magnetization that is lower than the value for the bulk material.

Since the magnetic moment of the single domains is proportional to the volume of the particle, it is very large compared to the individual magnetic moments of the ions of ordinary paramagnetic materials. As a result, the magnetic energy (even at moderate external magnetic fields) becomes comparable to the thermal energy. This implies that magnetic saturation is achieved at moderate fields and a high magnetic susceptibility is obtained for the particle system even at room temperature. When the relaxation time, a measure of the time for magnetization reversal, of the single domain system is less than the measuring time (the measuring time is a characteristic time for the used detection method) the particle system is called superparamagnetic [232]. The magnetization of the superparamagnetic single domain system can be calculated with the Langevin function averaged over a domain size distribution [233]. This model can be further modified for magnetic single domain interactions and magnetic anisotropy [234,235].

7.2. Néel relaxation

In a single domain particle system, where the particles are placed in a solid matrix such that the magnetic particles cannot physically rotate, the only change of the direction of the magnetic moment of the particle comes from an
internal rotation of the magnetic moment over an energy barrier whose size depends on the particular material. This type of magnetic relaxation is named Ne´el relaxation. The value of the energy barrier is depending on the material and size of the nanoparticle; along with the applied external magnetic field and the magnetic interactions between the nanoparticles that also changes the energy barrier. The magnetic moment can relax (or rotate) over the energy barrier during a specific time (the relaxation time) if there is enough thermal energy available. The magnetic relaxation time of magnetic nanoparticles can approximately be considered as the time for the magnetic moment of the nanoparticle to rotate from one easy axis direction to the other easy axis direction. The Ne´el relaxation time, \( t_N \), is dependent on the temperature, \( T \), the magnetic anisotropy, \( K \) (which is dependent on the material, size and the geometry of the nanoparticles) and the volume of the nanoparticles, \( V \). The Ne´el relaxation time, in zero magnetic field and for non-interacting nanoparticles, can be expressed by [236]:

\[
\tau_N = \tau_0 e^{(KV/k_B T)} \tag{5}
\]

where \( k_B \) is the Boltzmann constant and \( \tau_0 \) a characteristic relaxation time in the range of \( 10^{-9} \) and \( 10^{-12} \) s [234,237]. Since the Ne´el relaxation time depends exponentially on the energy barrier, the Ne´el relaxation time varies drastically with the size of the nanoparticles as can bee seen in Fig. 16. The Ne´el relaxation time changes from very low values of about nanoseconds up to several thousands of years when the size of the nanoparticles changes. This enables control of the relaxation time and magnetic properties of the nanoparticles and ultimately the design of magnetic nanoparticles for specific applications.

Table 2

<table>
<thead>
<tr>
<th>Material</th>
<th>( d_S ) (nm)</th>
<th>( d_B ) (nm)</th>
<th>( M_s ) (kA/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FeO₄ ³⁻</td>
<td>70</td>
<td>15</td>
<td>390</td>
</tr>
<tr>
<td>γ-Fe₂O₃</td>
<td>70</td>
<td>20</td>
<td>320</td>
</tr>
<tr>
<td>Fe</td>
<td>50</td>
<td>15</td>
<td>1500</td>
</tr>
<tr>
<td>Co</td>
<td>30</td>
<td>7</td>
<td>1300</td>
</tr>
<tr>
<td>CoFe₂O₄</td>
<td>90</td>
<td>10</td>
<td>380</td>
</tr>
<tr>
<td>FePt</td>
<td>50</td>
<td>5</td>
<td>840</td>
</tr>
</tbody>
</table>

The \( M_s \) values are given for small magnetic nanoparticles (which is usually smaller than the corresponding values for bulk values as explained in the text). The diameter \( d_B \) is explained in the text.

Fig. 16. The plot shows how the Ne´el relaxation time (blue circles) and Brownian relaxation time (green rectangles) varies with the diameter of the single domains. The parameters used in the plot are typical data of single domains of maghemite (γ-Fe₂O₃) at room temperature. A size dependence of the magnetic anisotropy has also been taken into account in this plot. The Brownian relaxation time is plotted for nanoparticles placed in water at room temperature with a hydrodynamic size of 20% larger than the actual domain size. The effective relaxation time (red line) is the relaxation time the particles will undergo when placed in, as in this case, water. Nanoparticles in a solid matrix will always only relax with the Ne´el relaxation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)
Magnetic nanoparticles with long Néel relaxation times can be used in data storage systems where it is crucial to have small regions of magnetic material with stable magnetic moment directions. The two directions of the magnetic moments of the magnetic nanoparticles enable the storage of binary data as zeros (0) and ones (1). The directions of the magnetic moment of the nanoparticles must be stable with time since otherwise the information would be lost. Today the research of using magnetic nanoparticles for information storage is under intensive development [238,239].

Magnetic nanoparticles can be subdivided into particles that are thermally blocked or superparamagnetic [232,237]. Thermally blocked particles have magnetic relaxation times that are longer than the measurement time (the characteristic time for the used detection method) and superparamagnetic particles that have magnetic relaxation times that are short compared to the measurement time. The size that divides the particles into superparamagnetic and thermally blocked particles is given in Table 2 (diameter $d_B$), and is dependent on material and temperature. If the nanoparticles are placed in a solid matrix the thermally blocked nanoparticles will exhibit remanence and coercivity while the superparamagnetic particles will not show any remanence and coercivity, e.g. the coercivity and remanence for a nanoparticle system depends on the magnetic relaxation compared to the measurement time. Coercivity is the field that brings the magnetization to zero value and the remanence is the residual magnetization of the particle system after being magnetized with an external magnetic field. Thermally blocked particles can then used in data storage systems and it is the direction of the remanence that yields the information and the coercivity gives the necessary magnetic field to write data to the information storage system.

7.3. Brownian relaxation

If the nanoparticles are placed in a carrier liquid the particles can rotate and another relaxation mechanism then becomes significant, Brownian relaxation. The Brownian relaxation time, $\tau_B$, is dependent on the viscosity of the liquid, $\eta$, the hydrodynamic volume, $V_h$, of the particles and the temperature. The Brownian relaxation time can be expressed as [240]:

$$\tau_B = \frac{3V_h\eta}{k_B T}$$  \hspace{1cm} (6)

The variation of Brownian relaxation time with nanoparticle size can also be seen in Fig. 16. The Brownian relaxation time depends also on the size of the nanoparticle but not as dramatically as the Néel relaxation time. The effective (or the total relaxation time) of the nanoparticles placed in the liquid is a combination of the Néel and Brownian relaxation time. The mechanism by which that the nanoparticle system relaxes is the mechanism with shortest relaxation time. By setting the Néel relaxation time equal to the Brownian relaxation time it is possible to determine a critical nanoparticle diameter that divides particles which relaxes with the Néel relaxation (internal relaxation of the particle magnetic moment) or with the Brownian relaxation (where the direction of the particle magnetic moment follows the rotation of the particle). For magnetic nanoparticles of maghemite at room temperature in water, the crossover between the two relaxation mechanisms occurs at 15 nm (see Fig. 16). Today, most of the magnetic nanoparticles in medicine and biomedicine use maghemite or magnetite nanoparticles equal to or below a diameter of about 10 nm [241]. The magnetic nanoparticles in these size ranges have short relaxation times, $\mu$s or less. In magnetic separation techniques in biomedical applications, the size of the total magnetic particles can be 1 $\mu$m or even larger. In this case, the particle is built up of several magnetic nanoparticles in the range of 10 nm in a polymer matrix. Using these small nanoparticles, with short relaxation times, agglomeration of particles after the particle system has been exposed to a magnetic field gradient (which is used in the magnetic separation process) is prevented. Magnetic nanoparticles with sizes of about 10 nm and below are also used as contrast agents in magnetic resonance imaging (MRI) [241].

7.4. Biosensor methods using magnetic nanoparticles

Magnetic nanoparticles in biosensor applications have been used for several years. Different detection techniques are being used in these applications, very sensitive superconducting quantum interference device (SQUID) sensors [242], giant magnetic resistance (GMR) sensors [243,244], flux-gate magnetometers [245] and induction techniques [246]. Both the Néel and the Brownian relaxation are used in these biosensor applications in order to detect the presence of different bio substances in a liquid. The methods that uses SQUID sensors and flux-gate magnetometers measures the magnetization decay of magnetic nanoparticles after they have bounded to a rigid surface (due to the
presence of a specific substance in the liquid) and magnetized by an external field. In this case, the Néel relaxation should be long enough which sets a lower limit of the size magnetic nanoparticles according to Fig. 16. In the case of the GMR technique, the magnetic nanoparticles is superparamagnetic with no remanence and consequently the particle system must be magnetized when the measurement is performed. This is achieved by an external magnetic field with a field direction not oriented in the sensitive direction of the GMR sensors.

In the biosensor method using induction techniques to detect the Brownian relaxation [246], the Néel relaxation time must be larger than the Brownian relaxation time. Then, the orientation of the magnetic moment of the particle changes with the same rate as the rotation rate of the particle itself, e.g. the Brownian relaxation is detected by magnetic detection. This put a lower limit of the sizes of the nanoparticles. From Fig. 16, the lowest size at room temperature for maghemite nanoparticles is about 15 nm. The magnetic nanoparticles in this biosensor method must then be larger than 15 nm. The method is based on measurements how the Brownian relaxation time changes when different substances in a liquid binds to the surface of the particles. The magnetic response, \( \chi(f) \), at different excitation frequencies, \( f \), of the applied field is measured and analyzed by a model including a distribution of Brownian relaxation times (mostly due to a spread in particle sizes) according to:

\[
\chi(f) = \int \chi(f, \tau_B) f(\tau_B) \, d\tau_B
\]  

where \( \chi(f, \tau_B) \) is the frequency dependent Debye response and \( f(\tau_B) \) the distribution function (a log-normal distribution is used in the model) of the Brownian relaxation time. The Brownian relaxation time can be translated to a size of the particle (Eq. (7)) and the size distribution of the particle system can be determined. When the excitation frequency approaches \( 1/(2\pi\tau_B) \) the imaginary part of the magnetic response is at its maximum. The graph in Fig. 17 shows the imaginary part of the magnetic response versus excitation frequency when different amounts of prostate specific antigen (PSA) are absorbed on the surface of the particles.

From Fig. 17, we can see that increasing the amount of PSA bounded to the particle surfaces gives a frequency shift of the magnetic response towards lower frequencies. This is due to that the hydrodynamic size of the particles increases and thereby the Brownian relaxation time increases, meaning a decrease in the corresponding Brownian relaxation frequency, \( 1/(2\pi\tau_B) \). From Fig. 17 (inset), we can also see that the shift in the Brownian relaxation frequency is saturated at higher PSA concentration. At this high PSA concentration in the liquid, the surfaces of the particle system are saturated with PSA. The dynamic range of the concentration detection is determined by the size of the particles (different surface areas) and the number of particles in the liquid. With this biosensor method it is possible to detect different kinds of substances in a liquid [246,247].

7.5. Conclusions

We have seen in this chapter that there are large numbers of different applications where magnetic nanoparticles are key components. Depending on the application different kinds of magnetic nanoparticles are used with different
magnetic relaxation times. As we have also seen, the magnetic properties of magnetic nanoparticles are closely related to the magnetic relaxation time. The magnetic relaxation can be tailored by changing the size of the nanoparticles or utilising different materials in the particles. Due to this and on the rapid development of nanotechnology, magnetic nanoparticles will be a great tool in a variety of industrial or medical applications, as well as an excellent model system for further scientific investigations.

8. Magnetic shape memory alloys

Magnetic shape memory (MSM) materials have received significant attention because of the large shape changes they produce when exposed to moderate magnetic fields (typically \(< 1 \text{ T}\)). The deformation of these materials can be as large as 10\% and operating frequencies well above 100 Hz. The materials return to their original shape when applying a spring force, or reversing the direction of the magnetic field by 90°. MSM materials are typically single crystalline metal alloys, which convert electrical power to mechanical power, and vice versa.

The first report on a large magnetically induced strain in a Dy single crystal is dated back in 1968 [248]. Liebermann and Graham [249] reported a giant 3.4\% reversible strain in the magnetically hard direction in a single crystal Dy subjected to a magnetic field of 100 kOe at 4.2 K. Detailed investigations revealed twinning to be the primary deformation mode. It was suggested that the driving force for the observed deformation is the lowering in the magnetostatic and magnetocrystalline anisotropy energies on twinning. However, the magnetic field-induced deformation received significant attention after similar strains were measured in Heusler type non-stoichiometric Ni$_2$MnGa alloys [250–255]. The exact mechanism of large reversible strains by the application of an external magnetic field is not thoroughly understood yet. It is proposed that this can be due to structural transformations in martensite phases [256–259], twin variant conversion and reorientation [260–263], or by the magnetic force generated due to non-uniform magnetic field, which can deform the martensite and induces twin variant rearrangement [264–266]. The magnetic shape memory effect essentially differs from magnetostriction, even though some effect of magnetostriction may contribute in the early stages of the deformation process. Finally, the effect of internal stresses, caused by structural transformation and alloy processing, in the magnetic field-induced deformation received little attention.

The martensite crystal, after transformation from cubic austenite, consists of the mixture of tetragonal martensite variants, having different c-axis orientation, separated by twin boundaries. In MSM materials, the twin boundaries are highly mobile. A simplified representation of the MSM deformation process is given in Fig. 18, where it is assumed that only two twin variants exist in the material. These variants have different magnetic and crystallographic orientations. When the MSM material is exposed to an external magnetic field the twins in a favourable orientation relative to the field direction “grow” at the expense of other twins. Magnetic field $H$ increases the amount of twin variants of “preferable” orientation, i.e. the twin boundaries move in the material. Steuwer et al. [265] demonstrated that large shape changes in the order of $10^{-2}$ cannot be achieved by a direct effect of the external magnetic field on twin variant rearrangement. On the other hand, the magnetic force, which is proportional to the magnetic field gradient, can be of a magnitude enough to cause twin variant rearrangement.

The mechanical stress needed to move the twin boundary is called twinning stress, $\sigma_{\text{tw}}$. This term is important when studying the magneto-mechanical properties of MSM materials. As the Ni–Mn–Ga MSM material has at room temperature a martensitic tetragonal crystal structure (lattice parameters $c/a < 1$, typically 0.94), the material exhibits a net shape change as a result of the change of the relative amounts of twin variants. That shape change equals the magnitude of the tetragonal distortion:

$$e_H = \frac{a}{c} - 1$$

Magnetocrystalline anisotropy has a key role in achieving large magnetic field-induced strains. The energy of a ferromagnetic crystal depends on the direction of the magnetization relative to the crystallographic axes. Tetragonal crystals show a uniaxial anisotropy. In this case, the energy density of the anisotropy is described as $E_a = K_{u1} \sin^2 \theta + K_{u2} \sin^4 \theta$, where $\theta$ is the angle between the $c$-axis (axis of tetragonality) and magnetization direction, $K_{u1}$ and $K_{u2}$ are anisotropy coefficients. A large positive $K_{u1}$ describes crystals with an “easy” axis of

24 Emmanouel Pagounis
magnetization along the c-axis. In the case of non-stoichiometric Ni$_2$MnGa alloys, the value of $K_{u1}$ was determined in the range from $1.3 \times 10^5$ to $2.48 \times 10^5$ J/m$^3$ [267–269] for different compositions. In the absence of an applied magnetic field, the magnetization vector lies along the easy magnetization axis. In the presence of a magnetic field, the magnetization rotates towards the direction of the field and magnetic energy of the variant increases. The maximum energy, which can be stored in the variants is given by the magnetic anisotropy energy, $E_a$, i.e. the energy originated from the rotation of magnetization vector from the easy magnetization axis to the direction of the field. We can explicitly equal magnetic anisotropy energy with elastic energy due to a twinning stress $s_{tw}$. When the magnetic field increases, the difference of magnetic energy given by the magnetic anisotropy energy $E_a$ exceeds the twinning stress energy ($s_{tw}H$) that is needed for twin boundary motion, and the twin boundaries are moving. Accordingly, important parameters for an existence of the MSM effect are high saturation magnetization, high magnetic anisotropy and low twinning stress.

8.1. Production and chemical composition

Magnetic shape memory materials are currently produced by conventional single crystal growth techniques, such as Bridgman [270,271] or Czochralski [272,273]. The details of the production process are rarely disclosed. After producing the single crystal bars the materials are homogenized at about 1000 °C for 24 h and ordered at 800 °C for another 20 h. Information about the effect of heat treatment on the crystal structure and the transformation temperatures is given in the literature [274–277]. The material is then oriented using X-ray techniques to produce the desired crystallographic structure for the MSM effect. Following the crystal orientation, the material is cut and thermomechanically treated. The key to obtaining high strains is to cut the samples so that the twin boundaries are aligned at 45° to the sample axis (when magnetic field is applied transverse to the bar).

Studies of MSM alloys focus on ways to produce materials in which short response times are combined with large reversible strains. Several Heusler alloys and intermetallic compounds like Fe–Pd, Fe–Pt, Co–Ni–Al, Fe–Ni–Co–Ti have been investigated in this connection. The most promising results to date have been achieved with the ferromagnetic Heusler alloy Ni$_2$MnGa, and more exactly the family of Ni$_{2+x+y}$Mn$_{1-x}$Ga$_{1-y}$. In this alloy, magnetically controlled strains of up to 9.5% have been measured [278].

The MSM effect in these alloys is observed when the material is in its martensite state, therefore its crystal structure is of prime importance. The Ni–Mn–Ga alloys show a variety of martensitic structures. The crystal structure of the martensite depends strongly on the chemical compositions [279–281]. When the possibility of the MSM effect is studied, the main interest is focused on the ferromagnetic twinned martensitic structures. The crystal structure of martensite affects the magnetic anisotropy [282,283] as well as the magneto-mechanical properties [284–286], and the chemical behaviour [287]. The martensitic structures, transformation temperatures, and the Curie point have been systematically studied as functions of the average number of valence electrons per atom ($e/\alpha$) [288–291]. It has been suggested [292] that the martensitic transformation temperature increases linearly with the $e/\alpha$ concentration. High transformation temperatures and Curie point are essential requirements in increasing the operating temperature range of MSM materials. Shifts in the Ni$_{2+x+y}$Mn$_{1-x}$Ga$_{1-y}$ composition can significantly affect the maximum operating
The temperature of the MSM material [291,292], i.e. the maximum austenite start \( (A_s) \) temperature. For the time being, the MSM effect has been observed at temperatures of up to 65 °C, however, for several applications higher operating temperatures are required.

The most important martensitic structures found in the Ni–Mn–Ga system are the five layered tetragonal martensite (5 M), the seven layered near-orthorhombic martensite (7 M), and the non-modulated tetragonal martensite (T). The 5 and 7 M structures have lattice parameters \( c/a < 1 \), while the T structure has \( c/a > 1 \). Among them the 5 M martensite has mostly been studied to date, and gives a theoretical maximum field-induced strain of 6% at room temperature. The 7 M structure provides a maximum strain of 10.7%, while the T structure has so far not demonstrated the MSM effect. The T structure on the other hand has shown a huge 20% mechanically induced strain [293], but the large twinning stress (>6 MPa) hinders the magnetic field-induced deformation. The exact crystal structure that appears during cooling depends on the composition of the alloy, the thermo-mechanical treatment, and the thermal stability of the different martensitic phases [294–296]. The T structure is the most stable one and the alloys transforming straight to the T martensite from the parent phase have typically transformation temperatures close or above the Curie point [297–299]. The 5 M structure transforms directly from the parent phase at lower temperatures close to the ambient one, while the 7 M martensite appears directly from the parent phase in a very narrow temperature range below the Curie point, and is highly compositional dependent [278,300–303]. Even though the 7 M structure provides the largest field-induced strain, the magnetic field-induced stress is low, and it needs a higher magnetic field for saturation. These, together with its very narrow composition and temperature range, limit its practical applications to date. In the following text most measurements and experimental results are carried out using MSM materials with 5 M martensitic structures.

Among other candidate systems for the MSM effect most attention received the Fe–Pd [304–307], Fe–Pt [308–310], Co–Ni–Al [311–316], Ni2MnAl [317–319] and Fe–Ni–Co–Ti [320,321]. In these alloys, measured field induced strains were demonstrated in the Fe3Pd alloy [305] and in Co–Ni alloys [322,323]. An advantage of Fe3Pd alloy is its excellent workability compared to Ni2MnGa. In addition, attempts to replace the Mn in Ni2MnGa with Fe or Co have been reported [324–326], as well as the addition of rare earth elements [327,328].

### 8.2. Magnetic and mechanical measurements

Magnetic and mechanical measurements in MSM materials can be carried out with the set-up described in ref. [329]. Since MSM materials are ferromagnetic and highly anisotropic their magnetization is high, it has a specific saturation value which depends on the direction of the applied field. Fig. 19 shows the magnetization curve of a Ni–Mn–Ga MSM material. In the first magnetization cycle (Fig. 19a), the curve exhibits some interesting features. Initially, the increase of magnetization is slow and nearly linear, suggesting that the magnetization process is controlled by magnetization rotation. At approximately 280 kA/m the magnetization suddenly rises and then levels off, indicating the martensitic \( c \)-axis, which is the ‘easy’ magnetization axis of the tetragonal martensitic phase, aligns itself with the external magnetic field. That peculiar shape of the magnetization curve is an indication of the MSM effect. On decreasing the field the magnetization remains at the saturation value towards the lower field. This results in large transient hysteresis at the first quadrant. This hysteresis occurs in the first cycle, while during the second cycle (Fig. 19b) the magnetization reaches its saturation value in a weak magnetic field. Further magnetization loops show fast saturation without appreciable hysteresis, indicating that the ‘easy’ axis of magnetization (i.e. the \( c \)-martensite axis) is now oriented along the field direction (Fig. 19b). A slight tilt of the magnetization curve is caused by a demagnetization field. Observed transient behaviour can be, however, easily restored. The saturation value of intrinsic induction \( B_s \) in Ni–Mn–Ga alloys was measured between 0.6 and 0.68 T, while the saturation field strength \( H_s \) varied between 520 and 720 kA/m [262,267,329]. The changes in the magnetization curve as a function of the strain in a Ni–Mn–Ga MSM material is described in ref. [329].

From the magnetization curve of the MSM material (Fig. 19a), it is possible to calculate the material’s energy product. The magnetic cycle energy density \( w_{e\,\text{MSM\,lim}} \) is determined by the area between “easy” and “hard” magnetic axis (anisotropy energy) [260]. Numeric integration shows that:

\[
\frac{w_{e\,\text{MSM\,lim}}}{H_{\text{MSM}} \, dB_{\text{MSM}}} = 190 \text{ kJ/m}^3
\] (9)
The value of 190 kJ/m³ for the maximum energy density of the MSM material is an order of magnitude higher than that reported for the highest energy density actuator material so far, Terfenol-D. If we further assume an operating frequency of \( f = 600 \text{ Hz} \), then the limit average electric power density of MSM-material \( p_{\text{e av lim}} \) has a value of:

\[
p_{\text{e av lim}} = f \cdot W_{\text{MSM lim}} = 600 \times 190 \times 10^3 = 114 \times 10^6 \text{ W/m}^3 = 114 \text{ MW/m}^3
\]

(10)

In Fig. 20, the field-induced MSM strain in a Ni–Mn–Ga alloy under a variety of mechanical loads is shown as a function of the magnetic field. The characteristics of this type of curves depend on the chemical composition of the alloy, its thermal and mechanical history, and the temperature. The measurement curves demonstrate the large work output (force \( \times \) stroke) capabilities of MSM materials. In the measurements, the direction of the magnetic field was perpendicular to the long sample axis and fixed constant compressive stress was applied along the sample axis, i.e. field and loading directions are normal to each other. The expansion is then measured in the stress direction. The contraction was observed in the field direction so that the volume of the sample remained constant.

It can be seen in Fig. 20 that the stroke of the MSM material depends on the external load and field. With increasing compressive stress the maximum of field-induced strain decreases and the field needed for straining the material increases. The maximum blocking force for MSM materials was measured between 3 and 9 MPa \([251,254,304]\). The hysteresis observed in the strain curves indicates that the MSM material possesses also excellent damping capabilities. Using the hysteretic properties in actuator design and biasing permanent magnets, it is possible to reduce the applied current, and, thus, power consumption, while keeping the material in a certain position.

Since MSM-phenomena can be obtained only if the force generated in the material by the magnetic field is larger than the force needed for reorientation of the single martensite variant to another, i.e. the twinning stress \( \sigma_{\text{tw}} \), this value becomes very important in characterizing the MSM material. A low twinning stress corresponds to higher work output (force \( \times \) stroke) of the MSM material, and also lower activating field. The later is very important in reducing the size and power consumption of an MSM actuating device. Twinning stress values of \(< 1 \text{ MPa} \) are sufficient for actuator...
operation, however, MSM samples with a twinning stress of <0.4 MPa will provide significant advantages compared to existing technologies in practical applications. Low twinning stress materials exhibit also smaller hysteresis. The importance of twinning stress can be seen, for example, in the efficiency \( \eta_{\text{MSM}} \) of the MSM material, i.e. the ratio of output mechanical energy to input electrical energy. In MSM materials having twinning stress of around 1 MPa (measured in Fig. 20) the efficiency of the MSM material is about 50\%, while if twinning stress reduces to 0.4 MPa the efficiency increases to 80\%, with the theoretical limit closing to 95\%. Efforts are currently in progress to reduce the twinning stress of the MSM materials. Fig. 21 presents theoretical calculations on the effect of the material’s twinning stress in the activation field and the efficiency \( \eta_{\text{MSM}} \).

The temperature dependence of the MSM effect is also of importance, however this effect has received little attention to date. Similar to other smart materials, such as piezo-ceramics and magnetostrictives, the properties and performance of Ni–Mn–Ga MSM materials are affected by the operating temperature. It was found that the lattice parameters, which affect the free strain of the material, are highly temperature dependant. For example, in 5 M martensites cooling causes a slight increase of the \( \alpha \)-axis and a large decrease in \( c \)-axis [330,331]. In addition, the twinning stress of Ni–Mn–Ga increases with the decreasing temperature [332]. The lowest twinning stress values for a given alloy composition are measured at temperatures close to the transformation ones. The saturation magnetization

![Fig. 20. Strain output and applied magnetic field at different pre-stress loads in a Ni–Mn–Ga MSM material.](image1)

![Fig. 21. Efficiency and threshold activation field as a function of the twinning stress in a Ni–Mn–Ga MSM material.](image2)
and the anisotropy energy of the Ni–Mn–Ga martensite increase when the temperature decreases. These effects have a direct impact on the performance of the MSM materials at different temperatures, as demonstrated in Fig. 22. For a given alloy composition, there is a specific lower temperature where the twinning stress $\sigma_{tw}$ is high enough and exceeds the magnetic field-induced stress $\sigma_{mag}$. Below this temperature no MSM effect is observed [331,332]. In 5M martensites this temperature is between $-160$ and $-20$ °C, depending on the chemical composition, processing and thermo-mechanical treatment. This critical temperature is important in actuator applications, while in the sensor applications there is no low temperature limit.

The basic equations for MSM materials are summarized in the following [267,332–335]:

$$\sigma = \pm \sigma_{MSM} + \sigma_{tw}(\varepsilon);$$  \hspace{1cm} (11)

$$\sigma_{MSM} = \sigma_{mag} \left[ 1 - \left( 1 - \frac{H}{H_s} \right)^2 \right]; H \leq H_s;$$  \hspace{1cm} (12)

$$\sigma_{MSM} = \sigma_{mag}; \quad H > H_s;$$  \hspace{1cm} (13)

$$\sigma_{mag} = \frac{K_u}{\varepsilon_0};$$  \hspace{1cm} (14)

$$K_u = \frac{B_{is}H_s}{2}$$  \hspace{1cm} (15)

$$H_s = \frac{B_{is}}{(\mu_{hd} - 1)\mu_0}$$  \hspace{1cm} (16)

$$\eta_{MSM} = \frac{\sigma_{mag} - 2\sigma_{tw}}{\sigma_{mag}}$$  \hspace{1cm} (17)

$$\varepsilon_H = \frac{a - c}{c};$$  \hspace{1cm} (18)

where $\sigma$ is the mechanical stress of the MSM-material; $\varepsilon$ is the relative strain of the MSM-material; $\sigma_{MSM}$ is the magnetic field-induced stress of the MSM-material; $\sigma_{tw}(\varepsilon)$ is the twinning stress of the MSM-material as function of the relative strain; $\sigma_{tw}$ is the twinning stress of MSM material; $\sigma_{mag}$ is the maximum anisotropic magnetic field-induced stress of the MSM-material; $H$ and $B$ are the magnetic field strength and flux density within the MSM-material; $H_s$ is the saturation magnetic field strength (i.e., field strength at which easy direction and hard direction magnetization curves cross each other); $K_u$ is the anisotropic energy density of the MSM-material (equal to $w_{e\,\text{MSM\,lim}}$); $\varepsilon_0 = \varepsilon_H$ is the free strain of the MSM-material; $\mu_{hd}$ is the relative permeability of the MSM-material along the magnetic hard axis; $\mu_0$ is the absolute permeability of the vacuum; $B_i = B - \mu_0 H$ is the intrinsic induction.
within MSM-material; $B_i$ is the saturated value of the intrinsic induction; $\eta_{\text{MSM}}$ is the coupling factor (efficiency) of MSM material; $a$ and $c$ are the geometric dimensions of the crystallographic axes.

Mechanical and magnetic properties of MSM materials are summarized in Table 3. In Table 4, the magnetic field-induced strains measured in various Ni–Mn–Ga samples are summarized.

### Table 3
Magnetic and mechanical properties of Ni–Mn–Ga MSM materials [340]

<table>
<thead>
<tr>
<th>Control field</th>
<th>Magnetic property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. strain $\varepsilon$ (µm/mm), linear</td>
<td>100</td>
</tr>
<tr>
<td>Work output, $\sigma_{\text{eq}} \times \varepsilon_f$ (MPa × µm/mm)$^a$</td>
<td>300</td>
</tr>
<tr>
<td>Young’s modulus (GPa)</td>
<td>7.7$^b$</td>
</tr>
<tr>
<td>Compressive strength (MPa)</td>
<td>700</td>
</tr>
<tr>
<td>Curie temperature (°C)</td>
<td>103$^c$</td>
</tr>
<tr>
<td>Max. operating temperature (°C)</td>
<td>65$^c$</td>
</tr>
<tr>
<td>Resistivity (Ω m)</td>
<td>$80 \times 10^{-8}$</td>
</tr>
<tr>
<td>Relative permittivity</td>
<td>NA</td>
</tr>
<tr>
<td>Relative permeability</td>
<td>1.5–40</td>
</tr>
<tr>
<td>Coupling factor (%)</td>
<td>75</td>
</tr>
<tr>
<td>Energy density (kJ/m$^3$)</td>
<td>190</td>
</tr>
<tr>
<td>Piezoelectric coefficient $d_{33}$, Piezomagnetic coefficient $d_{31}$</td>
<td>NA 250 nm/A</td>
</tr>
<tr>
<td>Field strength for max. strain</td>
<td>400 kA/m</td>
</tr>
</tbody>
</table>

$^a$ $\sigma_{\text{eq}}$: blocking stress (in MSM defined as the stress at which magnetic field-induced strain is 0.01$\varepsilon_f$), $\varepsilon_f$: maximum strain.

$^b$ Measured from the compressive curve when the material is in single variant state.

$^c$ Maximum temperature achieved with current 5 M Ni–Mn–Ga MSM material.

### Table 4
Measured field-induced strain, stress $\sigma_{\text{mag}}$, twinning stress $\sigma_{\text{tw}}$ and anisotropy energy $K_u$ for some Ni–Mn–Ga alloy compositions

<table>
<thead>
<tr>
<th>Composition (at %) Ni–Mn–Ga</th>
<th>Strain (%)</th>
<th>$\sigma_{\text{mag}}$ (MPa)</th>
<th>$\sigma_{\text{tw}}$ (MPa)</th>
<th>$K_u$ 10$^5$ × J/m$^3$</th>
<th>Ms (K)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>46.6–29.5–23.9</td>
<td>2.2</td>
<td>&gt;2</td>
<td>1.22</td>
<td>&gt;300</td>
<td>[341]</td>
<td></td>
</tr>
<tr>
<td>47.4–32.1–20.5</td>
<td>5.7</td>
<td>&gt;2</td>
<td>1.7</td>
<td>301</td>
<td>[342]</td>
<td></td>
</tr>
<tr>
<td>48–30–22</td>
<td>7.3</td>
<td>&gt;1.5</td>
<td>2.13</td>
<td>307</td>
<td>[269]</td>
<td></td>
</tr>
<tr>
<td>48–31–21</td>
<td>6</td>
<td>&gt;1.5</td>
<td>1.7</td>
<td>297</td>
<td>[343]</td>
<td></td>
</tr>
<tr>
<td>48.1–29.4–22.5</td>
<td>4.5</td>
<td>&gt;1.5</td>
<td>1.5</td>
<td>302</td>
<td>[344]</td>
<td></td>
</tr>
<tr>
<td>48.2–30.8–21</td>
<td>9.5</td>
<td>1.5</td>
<td>1.6</td>
<td>337</td>
<td>[283]</td>
<td></td>
</tr>
<tr>
<td>48.8–29.7–21.5$^a$</td>
<td>9.4</td>
<td>1.9</td>
<td>1.6</td>
<td>337</td>
<td>[345]</td>
<td></td>
</tr>
<tr>
<td>48.9–29.6–21.4</td>
<td>3.8</td>
<td>2.14</td>
<td>2</td>
<td>&gt;300</td>
<td>[346]</td>
<td></td>
</tr>
<tr>
<td>49.2–29.6–21.2</td>
<td>4.78</td>
<td>2</td>
<td>1.22</td>
<td>&gt;290</td>
<td>[347]</td>
<td></td>
</tr>
<tr>
<td>49.2–29.6–21.2</td>
<td>5.8</td>
<td>2.1</td>
<td>1.45</td>
<td>&gt;290</td>
<td>[233]</td>
<td></td>
</tr>
<tr>
<td>49.5–25.4–25.1</td>
<td>0.4</td>
<td>&gt;2</td>
<td>1.86</td>
<td>306</td>
<td>[234]</td>
<td></td>
</tr>
<tr>
<td>49.6–28.4–22</td>
<td>0.5–5</td>
<td>&gt;1</td>
<td>2.48</td>
<td>305</td>
<td>[269,348,349]</td>
<td></td>
</tr>
<tr>
<td>49.7–28.7–21.6</td>
<td>5.3</td>
<td>&gt;1</td>
<td>1.5</td>
<td>318</td>
<td>[350]</td>
<td></td>
</tr>
<tr>
<td>49.7–29–21.3</td>
<td>2.6$^b$</td>
<td>&gt;1</td>
<td>1.8</td>
<td>&gt;310</td>
<td>[355]</td>
<td></td>
</tr>
<tr>
<td>49.8–28.5–21.7</td>
<td>6</td>
<td>&gt;1</td>
<td>1.5</td>
<td>318</td>
<td>[253]</td>
<td></td>
</tr>
<tr>
<td>50.5–29.4–20.1$^c$</td>
<td>0.47</td>
<td>1.6</td>
<td>3.5</td>
<td>1.74</td>
<td>[350]</td>
<td></td>
</tr>
<tr>
<td>50.5–30.4–19.1$^c$</td>
<td>None</td>
<td>0.7</td>
<td>17–25</td>
<td>1.44</td>
<td>&gt;310</td>
<td>[350]</td>
</tr>
<tr>
<td>50.7–28.4–20.9</td>
<td>&gt;6</td>
<td>2.8</td>
<td>&gt;1</td>
<td>1.67</td>
<td>&gt;290</td>
<td>[351]</td>
</tr>
<tr>
<td>51.1–24–24.9</td>
<td>5.8</td>
<td>5.7</td>
<td>1–3</td>
<td>178</td>
<td>[352]</td>
<td></td>
</tr>
<tr>
<td>51.3–24–24.7</td>
<td>0.5–1.3</td>
<td>&gt;2</td>
<td>2.45</td>
<td>264</td>
<td>[251,268]</td>
<td></td>
</tr>
<tr>
<td>52–24–23.6</td>
<td>0.6–1.2</td>
<td>&gt;2</td>
<td>2.45</td>
<td>276</td>
<td>[351]</td>
<td></td>
</tr>
<tr>
<td>52.1–27.3–20.6$^c$</td>
<td>&lt;0.02</td>
<td>1.1</td>
<td>15</td>
<td>2.03</td>
<td>[283,346]</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ 7 M martensitic structure.

$^b$ At frequency 2 Hz and external load of 1.86 MPa.

$^c$ T martensitic structure.
8.3. Magnetic shape memory actuators

Magnetic shape memory materials have been developed primarily for actuator applications. The operation principle of an MSM actuator is schematically presented in Fig. 23. The actuator consists of two coils aligned symmetrically to the MSM element, a prestress mechanism and the appropriate ferromagnetic core (not shown in the drawing). Before applying the magnetic field, the MSM element is aligned with its short martensite crystallographic c-axis along the stick axis (i.e. the direction of pre-stress loading). When a magnetic field is applied the MSM element elongates in the direction perpendicular to the field. When the field is removed the pre-stress mechanism, usually a mechanical spring, contracts the MSM material to its original length. The operation is then continuously repeated. The electric and magnetic circuit of an MSM actuator is modelled in ref. [335].

The basic feature of an MSM actuator is that the magnetic field is applied perpendicular to the MSM material. In the topology presented in Fig. 23, the MSM element is located in the same magnetic circuit with ferromagnetic core and the magnetic field is generated with the coils. The air-gap should be kept as small as possible to eliminate power losses. In actuator design, it is also possible generate the necessary magnetic field using only one coil. The magnetic circuit to drive the MSM stick shall be designed to give field strength on the surface of the sample 300–500 kA/m. This corresponds to a flux density of about 0.6 T on the surface, and of about 1.3 T inside the MSM material. The difference between the external and internal field is due to the demagnetization effect observed in ferromagnetic materials [354]. Lower magnetic fields can be applied when the twinning stress of the MSM material is reduced (Fig. 21). This has a direct effect in reducing the size and power consumption of the actuator, leading to increased efficiency.

Biasing permanent magnets (PM) are often used to increase the field strength and reduce power consumption of the coils. In this case, the mechanical frequency of the device is the same as the electrical frequency at sinusoidal excitation. In practice, the permanent magnet generated dc field is half of the maximum magnetic field needed to drive the MSM material. Actuators with PM have usually a reduced size.

The pre-stress on the sample is usually between 0.5 and 1 MPa, a spring or disk spring can be used. Correct pre-stressing is crucial in optimising actuator operation, as it affects the actuator’s force and stroke capability (see also Fig. 20). Both the spring force and the load the actuator is working against should be taken into account in the actuator design. An optimal load to reach maximum magnetic-field-induced strain is about 1–1.5 MPa (Fig. 20); however, values of 2 MPa or higher can be reached with proper alloy development and low twinning stress materials [333,353].

The MSM element, the moving mass and the pre-stress spring are the basic components of the mechanical circuit of the MSM actuator. Since the actuator can work at a high frequency, the resonance frequency of the mechanical system is often reached. In case the actuator is used only at a specific frequency and mass the mechanical resonance frequency can also be used to increase the motion of the system.

The properties of the MSM actuator are strongly affected by the core type of the magnetic circuit. In particular, eddy currents have to be minimized in high frequency applications. Fig. 24 presents a curve, computed from field theory, which shows the field penetration inside the MSM material. It can be seen that, theoretically, with 2.5-mm thickness of the MSM stick the maximum operating frequency can be well above 1000 Hz. In practice, current device

---

Fig. 23. Basic structure of an MSM actuator.
work is often at lower frequencies due to the existence of eddy currents within the material. Improved magnetic circuit design and a better understanding of the properties of the material should enable higher operating frequencies. Laminated MSM materials can be also used when higher frequencies are required.

The electric circuit connected to the coil senses the construction in Fig. 23 as an inductor. The magnetic field strength inside the MSM element is roughly dependent only on the applied electric current. The voltage supplied to the coils has a delay before it affects the magnetic field inside the MSM element. Therefore, MSM actuators are best driven with current sources. If the actuator is driven with the voltage source, an additional delay in the operation may exist; as a result best performance is achieved when using a power supply with current controller.

When exposed to a magnetic field the MSM element generates a stress. Stress is a function of the magnetic field strength, which in turn is a function of the input current. Hence, when the current it controlled it is possible to control the MSM actuator’s acceleration. This relation affects the position control of the actuator. For position control applications, it is beneficial to have an intelligent control algorithm, which takes the relation between current and motion into account.

In Fig. 25, a response time of 0.7 ms is demonstrated for an MSM actuator, however, the shortest rise times to reach strokes of 5% of the actuating element measured 0.2 ms [286,336]. The operating speed of the MSM actuators is limited only by the eddy-currents in the core and the inertia of the moving parts of the actuator. The actuator’s acceleration depends also on the moving mass. The largest acceleration is achieved when the moving mass is smallest, theoretically even $10^5$ m/s$^2$ can be reached. In Fig. 25, acceleration of $5 \times 10^3$ m/s$^2$ was measured in an MSM actuator.
The acceleration of the MSM element affects the maximum operating frequency of the MSM actuator. To date, operating frequencies from dc up to the kHz range have been achieved with MSM actuators [286, 336, 337, 355]. Higher frequencies can be reached with more efficient magnetic circuit design and power electronics. The limiting factor is often the voltage supply’s ability to change current in the actuator. Apart from the acceleration, the speed of an MSM actuator is also an important design parameter. In an MSM material, a maximum speed of 1.5 m/s was measured [336] with the theoretical limit being the speed of sound (333 m/s).

It is useful to include an external position measurement for position control applications since the motion of the MSM actuator depends on the force the MSM element has to work against. A diagram of a control system for a MSM actuator can be seen in ref. [352]. The control system is designed for a linear motor’s MSM actuator, but it is generally usable for any MSM actuator used for position control. Ultra-high (nm) positioning resolution can be achieved, depending on the resolution of the position sensor used.

Hysteresis exists between the strain and the actuator input current, as well as between the strain and stress. Hysteresis is caused by the internal properties of the MSM material, such as twinning. The hysteresis of the MSM material (Fig. 20) has to be taken into account in the design of a specific MSM actuator application. The hysteresis dampens unwanted mechanical vibrations and higher harmonics of current and in that way eases the control of the application. In addition, it reduces vibrations and overshooting of the MSM actuator in rapid shape changes of the element. An inherent feature of the MSM actuators is that the stroke can remain in a certain position, even under external load, without using current. This can be achieved using the hysteresis of the MSM material and proper actuator design with biasing permanent magnets.

The magnetic field-induced force of an MSM actuator depends on the cross-sectional area of the MSM element, with a typical blocking stress of 3 MPa. The stroke in turn depends on the length of the MSM element, typically in actuator operation the MSM material produces up to 5% elongation. These affect the dimensions and performance of the actuator [334]. In actuator operation, MSM materials have demonstrated more than $5 \times 10^8$ cycles without deterioration of stroke [356]. These results need yet to be confirmed in practical actuator applications. Fatigue properties and crack propagation in MSM materials have recently been studied in more detail [338, 339]. Since the MSM shape change is relatively large, the motion can often be used directly without any mechanical amplification. Hence, the structure of the actuator can be simple and reliable, as backlash and wear problems are eliminated with this kind of construction.

Fig. 26 summarizes the performance of various actuator technologies [340]. As can be seen, the MSM actuators bridge the gap between “smart” and conventional actuator technologies. Compared to conventional actuator technologies, the magnetic shape memory bridges the gap between ‘smart’ and conventional actuator technologies.
technologies (electromagnetic, hydraulics and pneumatics) MSM actuators offer advantages in terms of fast response, reduced size, high operation frequencies, enhanced controllability, reliability and efficiency.

8.4. Magnetic shape memory sensors, thin films and composites

Like most active materials, MSM materials possess also the inverse effect, i.e. mechanical straining of the material causes changes in the surrounding magnetic field. Generally, magneto-elastic materials have been used as sensing elements in force, torque or displacement sensors [357]. The sensor application of MSM materials has not been studied as much as the actuator one, however, studies revealed that MSM materials can be used in applications such as force, position or acceleration sensors [329,333,335,358]. In these studies, it was demonstrated that the MSM material can also be used in direct power generation, as large voltage (100 V) signals were induced by fast compressing the material, as shown in Fig. 27. Furthermore, the voltage signal is directly proportional to the velocity of the compression. A great advantage of the sensor approach is that only low magnetic fields are needed, which simplifies the design process.

As the magnetization depends on the strain state of the MSM material, the most obvious application is in a position sensor. This can be implemented by measuring the magnetization [329] or properties related to it, like the inductance. On the other hand, the material’s shape affects the magnetic flux going through it and, due to Faraday’s law, the change in flux can be monitored through the induced voltages. Since the induced voltages are proportional to the time derivative of the flux change [335,358] this phenomenon can be used in a speed sensor.

An interesting area for MSM materials is thin films due the potential for applications in micro-electro-mechanical systems (MEMS) [359–367]. The great challenge here is to grow uniform, defect-free, single crystalline films. Various methods have been used, like ion-beam sputtering [359], epitaxial growth on GaAs [360], pulsed laser deposition [361], r.f.-sputtering [362] and laser-beam ablation [365].

Finally, an area where MSM materials are also finding potential applications is that of ribbons and composites [368,369,326,370,371]. In composite materials, MSM alloy in plate, fibre or powder form is combined with the polymer matrix. Accordingly, the selection of a suitable polymer matrix, which will minimize potential reactions and accommodate the strain of the MSM alloy, is of prime importance. MSM composites offer the advantage that less bulk material is needed and high frequencies can be achieved, while the polymer matrix can be used as the return spring force. In these applications, the damping capabilities of composite MSM materials have also received attention.
9. Shape memory thin films for smart actuators

Recent progress in research and fabrication of thin films of shape memory alloys (SMAs) has opened the opportunity to develop novel actuators and systems, which are capable to perform smart functions by responding to their thermal, mechanical and magnetic environment in a controllable way [372]. The fabrication and material properties of SMA thin films have been described in a number of books and review articles, see, e.g. [372–376]. Until now, demonstrator devices have been developed for a broad range of applications, e.g., in the fields of microfluidics, robotics and information technology. The following section covers the aspects of design, modeling, fabrication and performance of selected examples of smart actuators of SMA thin films. New development trends in this exciting interdisciplinary field of materials research and engineering are highlighted.

9.1. Microfluidic valves using SMA thin films

The development of a SMA thin film microvalve has been pioneered by Johnson based on silicon technology using a monolithic integration approach [377–379]. Fig. 28 shows an exploded view of the microvalve structure. It consists of three silicon layers. The top layer is micromachined to form a spring. The middle layer contains a poppet supported by ribbons of NiTi. The valve seat is located in the bottom layer. NiTi ribbons and Si spring form an actuation system with passive biasing mechanism. At room temperature, the microspring forces the poppet against the valve seat to close the valve while the ribbons are stretched. In this condition, the force of microspring determines the maximum controllable pressure difference. Direct heating of the ribbons with electrical current causes the ribbons to contract pulling the poppet back into the plane of the middle Si layer.

Another technological approach has been to hybrid integrate micromachined SMA thin films, which allows the flexible combination with various materials such as thermoplastic substrates [380]. A schematic of such a microvalve is shown in Fig. 29. The main components are a plastic housing with an integrated valve seat, a polyimide membrane, a spacer, a SMA microactuator and a cover. The SMA microactuator consists of a circular array of microbridges.

Fabrication of the SMA microactuator is performed by chemical micromachining of a cold-rolled TiNi sheet of thickness $d = 20 \mu m$. In the unheated condition, a pressure difference between inlet and outlet causes a deflection of the membrane and the SMA microactuator. Thus, the valve is in a normally open condition. By application of an electrical heating current, work is generated, which is used for the control of fluid flow and pressure difference between inlet and outlet. The maximum backpressure for closing the valve is adjusted by the spacer thickness, which prestrains the SMA microactuator. Pressure compensation between the valve chambers above and below the membrane may be achieved by lithographically microfabricated membrane holes.

---

25 Manfred Kohl
For an optimum use of the shape memory effect with respect to work output and fatigue, several design criteria have been developed [372]. In short, it is essential to design the geometry and thermal connection to the environment such that homogenous profiles of stress and temperature of the mechanically active parts are created upon loading. Several valve prototypes have been developed, which largely fulfil these criteria [380–384].

Some major specifications of the SMA microvalves are summarized in Table 5. These specifications highlight the major advantages and drawbacks of using SMA thin films. The NiTi microvalves are capable to generate large work outputs and power densities in the order of 50 mJ/g and 1 W/g, respectively, which can hardly be reached by other actuation principles. By keeping the maximum strain below 1%, large cycling times well above $10^6$ are achieved. Due to thermal actuation, however, the frequency for complete actuation cycles is limited to about 10 Hz for the present designs (Table 5).

The dynamics of TiNi microvalves and temperature range of operation can be improved by using alternative SMA materials with phase transformation temperatures well above room temperature [383]. Alternatives are, for instance, sputtered TiNiPd thin films, whose phase transformation temperatures can be raised by several hundred degrees by increasing the Pd content [380,385,386]. Fig. 30 shows time-resolved gas flow characteristics of a NiTiPd microvalve for a pressure difference of 30 kPa at various ambient temperatures between room temperature and 120 °C. The microvalve is operated by applying periodic heating pulses of 1 ms duration. The heating power is adjusted to close the microvalve within about 10 ms. When no extra heating power is supplied after closing the valve, a minimum opening time of 22 ms is observed at 25 °C. Thus, the maximum operation frequency at room temperature is improved by a factor of about 3.5 compared to NiTi microvalves of similar layout.

Fig. 31 shows typical characteristics of gas flow of a NiTiPd microvalve and corresponding electrical resistance of the microactuator for a fixed pressure difference of 70 kPa. The flow characteristics of the microvalve

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specifications of selected SMA thin film microvalves</td>
</tr>
<tr>
<td>Materials</td>
</tr>
<tr>
<td>NiTi, silicon</td>
</tr>
<tr>
<td>NiTi, PMMA</td>
</tr>
<tr>
<td>NiTiPd, PES</td>
</tr>
</tbody>
</table>

The abbreviations PMMA and PES denote polymethylmethacrylate and polyethersulfone, respectively.
are determined in a thermostat at various ambient temperatures between room temperature and 120 °C using nitrogen gas. The resistance characteristics indicate that the change of gas flow is caused by the martensitic phase transformation. At 120 °C, the critical heating power required to close the valve is about 80 mW. With decreasing temperature, the power requirements increase reaching about 260 mW at room temperature. At 70 kPa, the maximum leakage rate is less than 0.1%. The specifications of the NiTiPd microvalve are summarized in Table 5.

The microvalve developments demonstrate the significant potential of SMA thin film microactuators. So far, SMA microvalves have only been fabricated in small scale production. For a broad introduction into the market, further developments in the fields of materials research and fabrication technology will be required in order to extend the application range and to minimize fabrication costs. For instance, suitable packaging and interconnection technologies have to be developed on the wafer level using parallel fabrication processes to mass-fabricate SMA microvalves. Apart from competitive stand-alone devices, current developments concentrate on microvalves for modular integration in fluidic systems of higher complexity.

---

**Fig. 30.** Time-resolved gas flow during closing and opening of a TiNiPd microvalve for a pressure difference of 30 kPa at various operation temperatures [386].

**Fig. 31.** Gas flow and electrical resistance characteristics of a TiNiPd microvalve for a pressure difference of 70 kPa at various operation temperatures as indicated [386].
9.2. Robotic devices using SMA thin film composites

Sputter deposited thin SMA films are also in use as composites in combination with a substrate. In this case, the substrate works as a bias spring and enables the SMA actuator to show a two-way behavior. Actuator motions with large deflections can, for instance, be generated by SMA thin films deposited onto metallic foils [387–389]. The motion occurs when film stresses induced by annealing treatment are drastically released by the shape memory film upon the martensitic transformation. Since these composites provide their work right after deposition and annealing without training they are promising candidates for microactuators.

By means of the combination of shape memory composites with additional polymer films, also bistable actuators can be realized [390]. Fig. 32 illustrates the functional principle of such a bistable shape memory actuator. The temperature-dependent stiffness of the polymer is superimposed to the mechanical hysteresis of the shape memory composite, which shows a two way behavior. The essential precondition of bistable behavior is that the glass transition of the thermoelastic polymer takes place within the hysteresis of the shape memory alloy. Recent investigations indicate that Ti$_{39}$Hf$_{15}$Ni$_{46}$/Mo composites can be used in combination with a Lucryl polymer to produce an actuator that has bistable behaviour [390].

Upon heating, first, the polymer becomes soft due to its glass transition at $T_g$ before the SMA starts to undergo the reverse martensitic transformation from martensite to austenite at the austenite start temperature $A_s$. Due to the low stiffness of the polymer above $T_g$, the actuator is able to move and since the SMA composite is deflected in the austenitic state the total polymer-SMA composite is bent. Upon cooling, first, the polymer becomes stiff below its glass transition temperature $T_g$ before the SMA starts to undergo the martensitic transformation at $M_s$. If the thickness of the polymer is high enough, the stiffness of the polymer below $T_g$ is sufficient to keep the SMA-composite in the curved state even after the transformation of the SMA to the martensite. Therefore, the actuator remains deformed. To switch the bistable actuator to the flat shape, the actuator has to be heated to a temperature between $T_g$ and $A_s$. Above the glass transition $T_g$, the polymer becomes soft and therefore releases the shape memory composite, which is still martensitic and tends to go to the flat shape as long as the temperature is below $A_s$. Upon cooling, the flat shape is fixed again by the polymer.

Fig. 33 shows a sketch of a microgripper using this mechanism. The gripper can be fabricated by magnetron sputtering and other MEMS-compatible methods on a silicon wafer.

Fig. 34 demonstrates that the polymer’s stiffness is strong enough to fix the actuator in the bent state (Fig. 34a) and to allow it to reach the flat state after heating to a temperature between $T_g$ and $A_s$ (Fig. 34b). The thickness of the polymer layers surrounding the metal is 75 $\mu$m and the thickness of the SMA composite is 35 $\mu$m (25 $\mu$m Mo + 10 $\mu$m SMA) in this case.

The transformation behavior of an actuator consisting of two neighbouring SMA thin film composites is automatically phase-coupled by the variation of temperature. Thus, complex motions can be easily controlled without

![Fig. 32. Stiffness vs. temperature of a thermoelastic polymer and transformation hysteresis of a SMA composite, required for bistability. Arrows indicate the heat exposures to change the curvature of the actuator [390].](image-url)
the need of additional electronics [391]. If the small hysteresis of a certain SMA is embedded within the broad hysteresis of the second SMA, a special transformation sequence can be provided. This mechanism enables the actuator to create a wave-like behavior which is very promising for the motion of microrobots as illustrated by Figs. 35 and 36. By deposition of the SMA thin films on opposite sides of the substrate, neighbouring SMA thin film composites are created, which work in opposite directions. Fig. 35 shows that, in this case, the path of deflection upon heating is completely different from the path upon cooling. Fig. 36 illustrates, how this concept can be used for locomotion of a robot. The cross-section of a robot leg is drawn as a superposition of the four different stages of a complete motion cycle. The SMA B with the broad hysteresis is placed at the upper part and SMA A at the lower part of the actuator. Every motion step corresponds to a different combination of austenite and martensite of the two SMAs. The stages are marked by the four possible phases. At the beginning, both alloys are martensitic and the leg is straight. By means of a single heat pulse, even legs with different initial positions can perform a complete motion cycle since a phase-coupling between these legs can be achieved directly.

Fig. 33. Principle of a bistable microgripper fabricated by magnetron sputtering [390].

Fig. 34. Bistable actuator (a) after 160 °C heat exposure in the bent state and (b) after 90 °C heat exposure in the flat shape, respectively. Photos taken at room temperature [390].

Fig. 35. Deflection-temperature measurement of a Ti$_52$Ni$_{34}$Cu$_{14}$/Mo/Ti$_{47}$Hf$_{11}$Ni$_{42}$ cantilever [391].
9.3. Microactuators of ferromagnetic SMA thin films for information technology

In the last decade, ferromagnetic shape memory alloys (FSMAs) experienced a widespread research boom as they exhibit a unique combination of high energy densities, shape memory and ferromagnetic properties [392]. Of particular interest have been the Heusler ferromagnetic shape memory alloys Ni–Mn–Ga. Recent investigations on bulk NiMnGa single crystals revealed the giant magnetostrain effect showing deformations as large as 10%, which is based on the magnetic field-induced reorientation of martensite variants [278,393]. In order to exploit these materials for microsystems applications considerable efforts have been undertaken recently to fabricate FSMA thin films [394–399]. The reader is also referred to Section 8 of this review.

The martensitic transformation temperature \( T_0 \) of Ni-Mn-Ga thin films can be tuned by the chemical composition in a large range from cryogenic temperatures up to temperatures above the Curie temperature \( T_C \), which is much less composition-dependent [400]. Thus, an optimum chemical composition with \( T_0 \approx T_C \) can be adjusted, where the material shows ferromagnetic thermoelastic martensites at room temperature and a mixed austenitic and ferromagnetic transformation in a relatively small temperature interval. This behaviour is of particular interest as the shape memory effect and the ferromagnetic transition can be induced simultaneously by a relatively small temperature change.

For FSMA thin films, a novel actuation mechanism has been developed, which makes use of both the ferromagnetic transition and the martensitic transformation [400]. The mechanism is illustrated in Fig. 37 for a Ni–Mn–Ga bending actuator placed in the inhomogeneous magnetic field of a miniature permanent magnet. Depending on the temperature of the microactuator, either magnetic or shape recovery forces occur in opposite directions, while the corresponding biasing forces remain small. Thus, an almost perfect antagonism can be realized in a single component part. By applying an alternating electrical current, a periodic oscillation of the beam can be excited. This motion can be used to control the deflection of a micromirror attached to the front end of the actuator. For instance, a microscanner prototype has been developed based on the actuation mechanism, see Fig. 38 [401]. The overall dimensions are 7 mm × 2 mm × 5 mm.

![Fig. 36. Superposition of the cross-section of a robot leg at different phases of a motion cycle [391].](image)

![Fig. 37. Operation principle of the FSMA microactuator. Legend: N,S, north and south pole of a permanent magnet; \( F_{\text{mag}} \), magnetic force; \( F_{\text{SMA}} \), shape recovery force; \( F_{\text{mart}} \), force in martensitic condition; \( \Delta \alpha \), mechanical scanning angle.](image)
Fig. 39 shows experimentally determined deflection angles of the microactuator in vertical direction $\Delta \alpha$ as a function of heating power in stationary condition [401]. For a heating power, lower than 40 mW, the double-beams remain nearby the magnet. At about 50 mW, the maximum temperature of the actuator reaches the Curie temperature causing a decrease of the deflection angle. This motion is supported by the shape recovery force for a heating power above 70 mW, where already a sufficient fraction of the double-beams is transformed to austenite. Above 100 mW, the flat position is reached.

A typical frequency characteristic is depicted in Fig. 40 with the optical scanning angle being twice the deflection angle of the FSMA microactuator $D_{\alpha}$ [402]. A pulsed heating signal is applied at a fixed duty cycle of 1:6 and fixed frequency steps of 2 Hz. For each frequency, the driving power is optimized. Surprisingly, it turns out that two optimum power regimes denoted as $P_{\text{opt1}}$ and $P_{\text{opt2}}$ exist, which give rise to large scanning angles. Below a critical frequency of about 120 Hz, the scanning angle shows several power-dependent maxima. These features occur due to the phase difference between heating pulses and eigenoscillations, which depend on the heating power [403]. By increasing the driving frequency above the critical frequency, the scanning angle sharply decreases independent of the heating power. This low-pass behavior can be attributed to the dynamics of heat transfer between the double-beam and its environment. The relatively short time constants for the given dimensions are mainly due to the high phase transformation temperatures and the forced convective cooling during scanning motion. Above about 150 Hz, a broad resonance occurs.

These results demonstrate that large scanning angles up to $60^\circ$ are generated for the given size of the FSMA microactuator. Furthermore, the operation frequency is not fixed but can be tuned in a wide range, which differs from
previous developments of microscanners working mostly at resonance [404]. The driving voltage remains below 1.5 V in all cases. Recent developments include extensions of design and technology to generate two-dimensional scanning patterns [405] and to combine the microscanner with a time-of-flight measurement setup for three-dimensional scanning of objects in the environment [402]. On main advantage of such systems is the small size and low moving masses allowing mobile applications with low sensitivity to vibrations and shock.

Various investigations have been undertaken to exploit the giant magnetostrain effect in FSMA thin films. Recent investigations on polycrystalline Ni–Mn–Ga thin films with a fiber texture show that prestrained films show stress-induced prealignment and magnetic field-induced reorientation of martensite variants [406]. However, the effects are relatively small, below 0.1%. Considerable improvements are expected for highly oriented and single crystalline FSMA thin films. First results on ex situ Ni–Mn–Ga thin films also show magnetic field-induced reorientation qualitatively [399,407]. It is expected that functional FSMA thin films with giant magnetostrain effect will be available in the near future opening up the development of novel microdevices with magnetic shape memory.

9.4. Conclusions

This review demonstrates that smart actuators fabricated from SMA thin films allow new solutions with remarkable performance reflecting the recent advances in the fields of material development, actuator design and modeling, micromachining, and system integration. Some of the presented microactuators are competitive components, which, however, have been fabricated in small quantities so far. Due to their high energy densities and multifunctional properties, they have a large potential for innovation and may contribute significantly to future developments in microsystems technology.

10. Shape memory materials

Shape-memory materials (SMM) are stimuli-responsive materials. They have the capability of changing their shape upon application of an external stimulus such as temperature or magnetic field. A change in shape caused by a change in temperature is called the thermally induced shape-memory effect. There are two main groups of SMM, metallic shape memory alloys (SMA) and shape memory polymers (SPM).

10.1. Shape memory alloys

SMAs are characterized by reverse phase transformation between two different crystal structures (phases). At temperatures lower than a predefined temperature $M_f$ (martensite finished temperature), the SMA exists in the martensite phase (Fig. 41). In the martensite phase, the SMA material is soft and ductile and it can be easily formed into any arbitrary shape with deformation up to 8%. When martensitic SMA is heated over $A_s$, (austenite start

Fig. 40. Scanning angle vs. frequency. Optimum power regimes are indicated by $P_{opt1}$ and $P_{opt2}$ [402].
temperature) it begins to transform into the **austenite** or **parent** phase. The transformation is finished at **austenite finish temperature** ($A_f$). During martensite → austenite transformation, not-constrained SMA material becomes mechanically hard and returns to the pre-deformed shape. After that, if the temperature falls again under the $M_f$ value, the SMA goes into reverse, austenite to martensite transformation. In reverse transformation, the specimen's shape remains unchanged until an external force is applied. With reference specifically to NiTi alloys, if the deformation of specimen in the martensite phase is modest, i.e. less than 5%, the process is fully reversible for millions of cycles. The transition temperatures of a SMA sample $M_s, M_f, A_s$ and $A_f$ are set by metallurgical composition as well as with the thermomechanical treatment history of the sample. Readers are referred to ref. [408] for more details.

If the SMA sample is fixed and clamped, that is to say it is prevented from returning to the shape it had in parent phase, the SMA can generate stress up to 800 MPa during transformation cycle. If the SMA material is used in an actuator, it can produce the specific work up to 5 J/g. This is the most energy-dense actuator for MST applications when compared with other actuator materials (electrostatic, electromagnetic, piezoelectric, etc.), as illustrated in Fig. 41b.

The ability of SMA materials to absorb the heat energy from their surroundings and convert it into the mechanical work during the martensite → austenite transformation and the reverse process, to dissipate the heat energy during deformation, can classify the SMA materials as “thermodynamics heat engines”. In the future, these characteristics could be applied to create a MEMS thermal engine. Theoretically, a SMA heat engine, using temperature difference from 0 to 100 °C, has the Carnot efficiency near 20%, close to that of gas turbines with efficiency of 35% [409] These attributes make the SMA useful for many applications beyond the well-known memory effect.

In the martensitic phase, SMA has excellent mechanical damping properties due to the energy absorption characteristics of the martensite structure. In addition, the martensitic form of SMA has an extraordinary fatigue resistance [410]. The mechanical energy of deformation is internally transformed to heat via fully reversible structural transformations. This property is known as “superplastic” behavior. The original parent shape can be recovered at any time simply by heating the sample above its transformation temperature. A thorough analysis of the damping properties of SMAs can be found in ref. [410].

In the austenitic (parent) phase, at temperatures slightly above the transformation temperature, the SMA's show another unusual effect, called “superelastic” or “pseudoelastic” behavior. For large deformations, the applied stress can induce the formation of martensite above the $M_s$ temperature (stress-induced transformation). During deformation the external force remains practically constant as the mechanical work is consumed for the phase transformation. Once a deformation of $\sim 8\%$ is reached, all the available austenite is transformed into the martensite and upon further deformation SMA responds as an ordinary elastic material. Since the stress-induced martensite has been formed above its normal temperature, the martensite reverts immediately to undeformed austenite as soon as the stress is removed. It should be noted here that all of the described effects can occur in the same specimen, depending on the temperature: under the $M_s$ the SMA has superplasticity behavior prevails, between $M_s$ and $A_f$ the material exhibits the memory effect.
and over $A_t$ behaviour is superelastic. These exciting features of the SMA materials sometime represent the serious drawback: without the serious consideration of all properties of the SMA materials, manufacturing and the learning procedures one may expect more from SMA, than it can offer.

The two major SMA systems that have significance today are the Ni–Ti alloys and the copper-base alloys [412]. Of these, TiNi is better suited for use as a microactuator. The Ni–Ti alloys have greater shape memory strain (up to 8%, 4–5% for copper-based alloys), have higher ductility and more stable transformation temperatures. Further, Ni–Ti alloys have excellent corrosion resistance, compared to the copper-base alloys; have better biocompatibility and the ability to be electrically heated for shape recovery [412].

The newly developed FeMnSiCrNi shape memory stainless steels (SMSS) have attracted substantial interest because of their good shape memory effect, mechanical properties, machinability, weldability and corrosion resistance [413]. The Ni–Ti has proven to be the most flexible and beneficial alloy for the micro-actuator application. One property makes Ni–Ti unique among other actuators for MEMS application; its power density rises with mass reduction, up to the nanometer range (Fig. 41b and Table 6).

The phase transformation of the Ni–Ti alloy was described in detail in 1965, when the term ‘memory’ was first used to describe the shape recovery behavior [415]. During the past four decades, the most of the complex behavior of Ni–Ti alloys was discovered, both practically and theoretically. The modern and comprehensive analysis of the Ni–Ti alloys can be found in ref. [408]. The most of knowledge about the Ni–Ti alloys is connected to the alloy in a bulk form, but many facts can be also applied to the thin film form.

10.2. Micro-scale applications of SMA

When one reviews the “new SMA materials for microsystems technology (MST) actuators”, the issue can be divided in two distinct fields. These are the application of the well known Ni–Ti alloys in the form of thin films, and also application and development of new alloys not previously used in micro-scale actuators.

Considering the application of Ni–Ti, MST brings the new challenges in the alloy manufacturing processes. The transition temperature of the NiTi alloys can be extremely sensitive to the composition. For Ni-rich NiTi compositions, a shift from the nominal composition of only 1% has been shown to change the transition temperature by 150°C [408]. Attaining the required composition of Ni–Ti alloy is a difficult task using thin film deposition techniques. In sputtering and laser ablation, the two main manufacturing processes, both have tolerance margins that are larger than those ideally required.

Sputtering deposition remains the principal manufacturing process for the Ti–Ni thin films. The process is MEMS compatible, which includes patterning and etching processes. The target is usually the Ti–Ni alloy; however, the composition of the sputtered film does not match the composition of the target. The most common solution to this problem is use of the dual phase target NiTi + Ti, or use of multiple sputtering guns. A comprehensive review of manufacturing processes can be found in refs. [416–440].

Considering the use of new SMA materials, most of the materials discovered in the past 40 years have no widespread application. The origin of this situation is complex: the alloy component cost, absence of a notable advantage over Ni–Ti alloy, manufacturing complexity, etc. For MST applications, the cost of the material is not the primary property, enabling materials previously ignored due to their high cost to be considered. Table 7 list some of the exotic SMA alloys [439]. The decisive answer about the use and development new shape memory materials for MST applications should be matter for future examination.

In principle, SMA actuators can be manufactured by simple deposition of the active material and they therefore are potentially easier to fabricate than electromagnetic or electrostatic actuators. In practice, however, as discussed

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Summary of the basic features of the SMA in micro-actuators [414]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>Disadvantages</td>
</tr>
<tr>
<td>High power/weight (or power/volume) ratio, increased with mass reduction Smooth, noiseless and friction free operation Excellent corrosion resistance and biocompatibility Simple and compact actuator mechanics</td>
<td>Low energy efficiency:, max 10% Limited bandwidth due to heating and cooling restrictions: Degradation and fatigue under the strain &gt;5% Complex control. The SMA behavior is influenced by a large number of parameters</td>
</tr>
</tbody>
</table>
elsewhere in this review (Section 9) and in the cited literature there are some practical barriers to monolithic integration of SMA films. Two of the most significant issues are film adhesion to substrates and compositional variation due to diffusion during heat treatment.

Applications of SMA to MST and microactuators started in the early 1990s. Since this time, many MST devices with SMA actuation have been developed including microfluidic, microoptic, valves and pumps [442–456]. The SMA actuators are usually heated resistively, but ambient temperature variations are also applied. Fig. 42 shows an SMA actuator activated by solar energy. When the sun illuminates the structure, the SMA ribs bend and open the highly reflective surface to the sun (Fig. 43).

10.3. Shape memory polymers

Shape memory polymers (SMP) are second important group of shape memory materials. Like SMAs, they have the capability of changing their shape in response to an external stimulus. The most common is temperature and thermo-responsive shape memory polymers typically consist of two polymer components and two phases, one with a higher melting temperature than the other. The glass transition temperature \( T_g \), is the reference point where the higher temperature component starts to melt. When heated above \( T_g \), the SMP is soft and rubbery and it is easy to change the shape. When subsequently cooled below \( T_g \), it retains the given shape (shape fixing characteristic). When heated again above \( T_g \), the material autonomously returns to the original “parent” shape [458–460].

<table>
<thead>
<tr>
<th>Alloy</th>
<th>Composition</th>
<th>Transition temp. (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag–Cd</td>
<td>44/49 at.% Cd</td>
<td>from –90 to –50</td>
</tr>
<tr>
<td>Au–Cd</td>
<td>46.5/50 at.% Cd</td>
<td>30 to 100</td>
</tr>
<tr>
<td>Cu–Al–Ni</td>
<td>14/14.5 wt.% Al, 3/4.5 wt.% Ni</td>
<td>–140 to 100</td>
</tr>
<tr>
<td>Cu–Sn</td>
<td>approx. 15 at.% Sn</td>
<td>–120 to 30</td>
</tr>
<tr>
<td>Cu–Zn</td>
<td>38.5/41.5 wt.% Zn</td>
<td>–180 to –10</td>
</tr>
<tr>
<td>Cu–Zn–X (X = Si,Sn,Al)</td>
<td>a few wt.% of X</td>
<td>–180 to 200</td>
</tr>
<tr>
<td>In–Ti</td>
<td>18/23 at.% Ti</td>
<td>60 to 100</td>
</tr>
<tr>
<td>Ni–Al</td>
<td>36/38 at.% Al</td>
<td>–180 to 100</td>
</tr>
<tr>
<td>Ni–Ti</td>
<td>49/51 at.% Ni</td>
<td>–50 to 110</td>
</tr>
<tr>
<td>Fe–Pt</td>
<td>approx. 25 at.% Pt</td>
<td>Approx. –130</td>
</tr>
<tr>
<td>Mn–Cu</td>
<td>5/35 at.% Cu</td>
<td>–250 to 180</td>
</tr>
<tr>
<td>Fe–Mn–Si</td>
<td>32 wt.% Mn, 6 wt.% Si</td>
<td>–200 to 150</td>
</tr>
<tr>
<td>NbU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nb3Sn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3Si</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 42. SMA actuator activated with solar energy. When the sun illuminates the structure, the SMA ribs bend and open the highly reflective surface (4 cm diameter) to the sun. [457].
The physical of memory effect in polymers is quite different from that in the SMA. The polymeric phase that exists above \( T_g \) corresponds to the higher melting temperature component. This acts as the physical cross-link that is responsible for the permanent shape. The second component acts as a molecular switch and helps to freeze temporary shapes below the transition temperature, with either the glass transition temperature or the melting temperature serving as the transition/switching temperature. Chemical cross-links can be used in elastomeric shape memory polymers instead of physical cross-links. The \( T_g \) of a SMP can be set over a temperature range of several hundred degrees by control of composition and/or the degree of cross-linkage [460].

The shape-memory effect is in principle a behavior inherent to all polymers. However, polymers that exhibit a useful shape-memory effect must demonstrate a sharp transition temperature and a rubbery plateau, along with relatively large strain capacity without local material damage. Only a few polymers that satisfy these criteria are described in the literature [458]. The common SMP’s surpass the metallic shape memory alloys in their shape-memory properties of maximum strain. Table 8 is a comparison of SMA and SMP materials showing the high maximum strain of the SMP, but with lower stress generating capacity compared to SMAs due to their lower stiffness.

The first SMP introduced the 1960s was the polyethylene, which is covalently cross-linked by means of ionizing radiation. However, SMPs was recognized as potentially attractive materials for actuators in the 1990s [459]. The next SMP was copolymer polyurethane along with terephthalate, polystyrene, poly(1,4-butadiene), polyethylene oxide etc. A thorough review of the SMPs can be found in ref. [460].

The range of actuation stimuli for SMP is, in general, wider than it is for SMA, which most frequently operate through application of heat. SMP materials can be heat-actuated but they can sometimes also be stimulated with light [461] and with water [462]. A notable hybrid structure of SMA and a separate polymer matrix has been actuated using a magnetic field [463]. This feature of SMP opens up new opportunities for novel the MST actuation systems.

### 10.4. SMP applications in MST

Due to relatively easy manufacturing processes and simple programming procedures of shape-memory polymers, these materials represent a cheap and efficient alternative to SMA. SMP find the broad spectrum of application that covers an area from minimally invasive surgery, through high-performance textiles, up to self-repairing plastic components. Although SMP materials display attractive properties for the microactuators, SMP have only limited application in the MST technology to date. One of the few applications of the SMP in micro devices is shown in Fig. 44, which show optical images of an active microfluidic reservoir during operation. The shape change of the SMP leads to black fluid travelling along the reservoir. It is expected that SMP will have more important role for microactuators in the future.

### Table 8
Comparison of the main properties of the SMA and SMP

<table>
<thead>
<tr>
<th>Actuator type</th>
<th>Max. strain (%)</th>
<th>Max. stress (MPa)</th>
<th>Specific elastic energy density (J/g)</th>
<th>Elastic energy density (J/cm³)</th>
<th>Maximum efficiency (%)</th>
<th>Specific density (g/cm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NiTi</td>
<td>&gt;5</td>
<td>&gt;200</td>
<td>&gt;15</td>
<td>&gt;100</td>
<td>&lt;10</td>
<td>6.5</td>
</tr>
<tr>
<td>Polymer</td>
<td>100–400</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>&lt;10</td>
<td>1</td>
</tr>
</tbody>
</table>
10.5. Conclusion

To conclude, a summary of the main properties and the difference between metallic and polymer memory materials will be made. The metallic memory materials (SMA) can deliver high force and have higher energy density. In contrast, the polymer materials (SMP) have a maximum strain that is two orders in magnitude larger than that of metallic materials.

The manufacturing technologies for MST applications are also very different. The metallic materials, mostly the NiTi alloy are made by sputter deposition. The characteristics of the deposited film are extremely sensitive to the processing parameters. The polymer memory materials are often spin coated onto a substrate and polymerized by thermal annealing.

The actuation stimulus for metallic materials in MST applications is generally a temperature rise. Note that magnetic shape memory materials are discussed in a separate section. For polymeric materials, specific compositions exist that can be actuated with heat but also with light, magnetic fields, water or they can respond to chemical reagents. This versatility sets them apart and it offers new possibilities in microtechnology for the design of actuators or sensors.

11. Expandable microsphere composites

Expandable microspheres are small plastic particles, which expand when heat is applied. The microspheres consist of a gas tight thermoplastic shell, which encapsulates a small amount of a liquid hydrocarbon, e.g. isobutane or isopentane. The dramatic expansion of the microspheres when heated is due to the fact that the liquid hydrocarbon is transformed into a gas. The shell of the expandable microspheres is made of copolymer precursors (monomers), for example, vinylidene chloride and acrylonitrile together with methylmethacrylate. When the microspheres are heated the thermoplastic shell softens and the liquid hydrocarbon inside the shell transforms into a gas, which increases its pressure. This results in a dramatic irreversible expansion of the microspheres (typical diameter values: from 10 to

---

27 Björn Samel, Wouter van der Wijngaart
40 μm), with a corresponding dramatic decrease of the density (typical values: from 1000 to 30 g/l). When fully expanded, the volume of the microspheres increases more than 60 times (Fig. 45a). After cooling down, the microspheres do not shrink because the thermoplastic shell hardens again and thus, the volume increase is irreversible.

Expandable microspheres are manufactured by a number of industrial groups, including Expancel [465] and Sekisui [466]. All experiments described in this section were performed with Expancel® beads.

The expandable microspheres are commercially available with expansion temperatures in the range of 80–190 °C, see Fig. 45b [465]. At temperatures above the temperature at which the highest expansion is obtained the microspheres gradually collapse. The expandable microspheres can be used in contact with many chemicals, including solvents without negative effects on expansion. The available microsphere sizes range from 6 to 90 μm in diameter (weight average diameter). Expandable microspheres are today used in a wide variety of applications in the macro-scale. Some examples include printing inks used to form topographical patterns on paper, in reaction injection moulding to counteract the shrinkage of polyurethane, and in a method to decrease the density of plastics through a controlled and predictable foaming process [465]. Moreover, expandable microspheres have been used to produce highly porous ceramics with high uniformity of the cell size and shape [467] used in car protection applications, for improved anticorrosion resistance and increased acoustical insulation levels [468].

In 2002, expandable microspheres were introduced as a novel actuator material within microsystem technology. This novel microactuator material has especially proven its performance in the field of microfluidics. Five specific technologies for the system integration of the actuator have been proposed: (1) direct mixing of the microspheres in liquid, (2) surface immobilization of the microspheres by incorporation in photoresist or (3) through self-assembly on a chemically altered surface, (4) through incorporation in a paste and (5) through incorporation as a composite in a polymer matrix.

In this section, we will discuss these five methods and some of the applications realized with them, including single-shot valves and pumps.

11.1. Direct mixing of the microspheres in liquid

Perhaps the most straightforward approach to incorporate expandable microspheres into a microfluidic actuator system is to pre-mix the expandable microspheres with a liquid. A mechanical barrier in the microchannel confines the actuator beads to a predefined area on-chip. Heating the liquid, either locally via a built-in heater or by heating the entire system, triggers the expansion of the microspheres and accomplishes the microfluidic actuation.
This method was proven successful for both pumping and valving of liquid [469]. The pump and valve system can be realized with nearly identical microstructures. They consist of a channel that contains a mechanical barrier (filter) for the microbeads, thus localizing them in a predefined area on-chip. The exact layout of the microfluidic device determines whether such a system works as a pump or as a valve, as illustrated in Fig. 46.

Test structures consisting of silicon microchannels covered by a glass lid were evaluated as valves where each channel has openings on both sides. Pumps were configured by sealing one opening of the microchannel with UV curing epoxy (Epotek OG 198).

The width of the inlet channel is typically 100 or 200 \( \mu \text{m} \), the width of the outlet 20 or 50 \( \mu \text{m} \). The depth of both channels is 50 \( \mu \text{m} \).

Liquid handling in microfluidic channels has been demonstrated as shown in Fig. 47. Water in the nanolitre range was displaced in a microchannel against a counter pressure of 100 kPa (i.e., the microspheres act as a one-shot pump). Otherwise the flow of water in a microchannel was completely blocked (i.e., the microspheres act as a one-shot normally open valve).

11.2. Surface immobilization of the microspheres by incorporation in photoresist

The microspheres can also be patterned and immobilized on a surface. Two techniques were developed for this purpose. In a first approach, the microspheres can be mixed in photoresist, and subsequently patterned. Best results were obtained using a lift-of process in a two-step spin process, see Fig. 48.

Two methods for selective immobilization of expandable microspheres without the use of mechanical barriers have been studied for their application in microfluidics, including patterning by photolithography and microcontact printing combined with self-assembly based on surface chemistry.

The microspheres were deposited on the surface of silicon wafers by a process based on a positive photoresist from Shipley (Microdeposit S1813).

![Fig. 46. Two concepts for liquid handling by Expandable microspheres: (a) is a one-shot pump and (b) a normally open one-shot valve [469].](image)

![Fig. 47. Micrographs showing a bead pack before (a) and after (b) a 70 °C heat treatment at 100 kPa counter pressure in a 200 \( \mu \text{m} \) wide channel. The total expansion is 1.4 nl [469].](image)
Some of the microspheres remained on the exposed area even though the photoresist was dissolved as can be seen in Fig. 48a. The remaining microspheres could be removed by overdeveloping the wafer, however, the selectivity to the unexposed area is then drastically decreased. Therefore, a layer of pure S1813 photoresist was first deposited, to serve as a release layer for the microspheres in the exposed area. This enabled removal of all microspheres from an exposed area after a development of 1.25 min, see Fig. 48b. The expandable microspheres were found to be transparent to the UV light used during the exposure. Hence, the microspheres did not shadow the photoresist and no distortion of the pattern was observed.

Fig. 49 shows selectively deposited expandable microspheres on a silicon wafer, using the photoresist technique described above, before (a) and after expansion (b and c). The shown pattern is a test pattern consisting of 100 µm pitched stripes. The expansion of the microspheres increases the thickness of the microsphere layer and alters the pattern-features to some extent as depicted in Fig. 49c. It was found that the density of deposited microspheres, i.e. the
number of microspheres per area, can be influenced by varying the amount of microspheres per photoresist volume or by varying the spin speed. This technique can also be applied to form patches of photoresist/microspheres on structured substrates, i.e. in etched channels. The channels can then be sealed, using a low temperature bonding technique which does not activate the immobilized microspheres, e.g. a PDMS film.

11.3. Surface immobilization of the microspheres through self-assembly on a chemically altered surface

Alternatively, the microspheres can also be patterned using self-assembly based on surface chemistry [471] utilizing micro contact printing technology. Hydroxy-functionalized beads have previously been immobilized on silicon substrates by means of surface chemistry. The beads were thought to be covalently bound to the substrate via ester bond formation between the anhydride-functionalized surface and the hydroxy groups on the beads. This assumption was corroborated by the fact that the beads did not attach when the printed 4-(chlorosulfonyl) benzoic acid is not transformed into anhydride. The surface chemistry of the expandable microspheres is currently not known. However, the same methodology as for the hydroxy-functionalized beads was applied for immobilization of the expandable microspheres. In Fig. 50, a scanning electron microscope (SEM) photo of an immobilized monolayer of
(unexpanded) expandable microspheres is shown. No microspheres were found outside the printed anhydride pattern showing that the microspheres can be patterned using this methodology. In Fig. 51, a SEM photo of expanded microspheres is shown. The microspheres remained bound to the substrate during and after the expansion. The immobilized microspheres (both unexpanded and expanded) could withstand rigorous washing in water. Preliminary experiments show that it is also possible to immobilize monolayers of expandable microspheres in deep reactive ion etched channels.

In such an application the expandable microspheres might be used to coat the internal walls of microfluidic channels in order to enlarge the reactive surface area. It should also be possible to chemically modify the surface groups on the microspheres in a variety of ways for screening and analysis within the fields of biochemistry and organic chemistry.

11.4. Incorporation of the microspheres in a paste

For mesoscopic applications (mm-range to cm-range), the microspheres can be incorporated in a paste. Such "active paste" is suited for integration during device assembly. A successful actuator approach was shown by mixing of microspheres and glycerine. Fabrication and testing of a thermally actuated one-shot liquid dispenser as an uncomplicated, fully functional, low cost device for use in medical disposables was shown.

The one-shot dispenser was designed to meet the requirements of disposable drug delivery system. Hence, all device components are made out of low cost materials and adapted for processes with the potential for high volume batch fabrication. Fig. 52 shows a schematic view of the device. A small portion of an expandable paste containing Expancel® microspheres and Glycerine is heated from the bottom side. When the paste reaches a temperature of 70 °C
it gradually expands into the liquid container. The liquid, separated from the paste by a thin elastic membrane, is pushed out as the expanded paste fills the whole container. Since, the volume expansion of the paste is manyfold, the heat conduction in the paste gradually drops to a considerably lower level, preventing the dispensed liquid from extensive heating. Fig. 53 depicts the disassembled device before and after expansion.

In another application the dispenser was used within a microfluidic dye laser system to create a regenerating flow of Rhodamine dye. This specific application requires a low flow rate rather than delivery of a precise dose. Therefore, the device was operated to deliver at low rates by decreasing the supplied actuation power. The lowest measured flow rate was 1 μl/h. By adjusting the electrical power to the heater the dynamic range of the flow rate is 1–2400 μl/h.

11.5. Incorporation of the microspheres as a composite in a polymer matrix

For large scale integration (LSI) of microfluidic actuators, it was proposed [470] that expandable microspheres are mixed in an elastomeric polymer matrix prior to application to the system. Incorporation of the expandable microspheres (XB) into the silicone elastomer poly(dimethylsiloxane) (PDMS), results in a composite material (PDMS-XB) that has the merits of both PDMS and the expandable microspheres.

In order to investigate replication molding techniques a master for replica molding was photolithographically manufactured using the positive photoresist AZ-4562 from Clariant. Thirty five micrometers deep structures were fabricated using standard lithography equipment to evaluate whether the composite retains the characteristic of pure PDMS to replicate features of the master with high fidelity. In the liquid pre-polymer state the composite was poured onto the mold, degassed for 5 min in vacuum, cured and eventually peeled off the master. Fig. 54 shows the high-resolution imprint of the replica-molded structure in the novel composite. Apparently, the novel composite is well suited for replica molding techniques, which significantly increases the design freedom of microfluidic components.

The basic property of the composite was investigated by means of a fluidic application and evaluated whether the composite completely fills out voids upon heating. A thin film of the composite was spun on a wafer and covered with a plastic replicated diffuser micropump with a depth of 80 μm serving as a precisely defined volume. To simplify matters expansion was effectuated all-over using a hot plate. The reservoir volume was consequently filled out.

![Fig. 53. Photographs showing the paste (a) unexpanded (state before actuation), (b) expanded (state after actuation; liquid container removed for visualization) and (c) expansion without constraining liquid container results in a volume expansion of seven times the original volume.](image-url)
Subsequent removal of the rigid reservoir revealed the expanded composite showing both its expansion potential and its capability of filling out deep as well as small voids. The latter was verified by the fact that the height of the expanded composite conforms to the depth of the reservoir (Fig. 55).

Single-use microfluidic pumps were fabricated in a multilayer fashion and show the displacement of liquids in the nanoliter range [471]. Localized heating is achieved through resistive heating elements patterned on a printed circuit board (PCB). All photolithographical fabrication steps use high-resolution transparencies. The PDMS-XB composite is obtained by thoroughly mixing PDMS base, curing agent (typically at 10:1 (v/v) base:curing agent) and XB at a concentration of 250 mg/ml liquid PDMS pre-polymer. A layer of the PDMS-XB composite is spun onto the patterned PCB at 750 rpm and subsequently degassed in vacuum and crosslinked at room temperature. An additional thin layer of the liquid PDMS pre-polymer is spun on top of the cross-linked composite at 2500 rpm.

Microfluidic channels and reservoirs are formed in PDMS by replica molding. The master was fabricated using Shipley SJR 5740 photoresist on a silicon wafer to yield inverse structures. Subsequent heating to 200 °C makes the inverse structures reflow which results in structures that are curved at the edges. Liquid PDMS pre-polymer is spun onto the master at 500 rpm, cured at 110 °C for 20 min and eventually peeled off the master.

The molded PDMS channels and reservoirs were bonded to the previously fabricated stack of PDMS/PDMS-XB/PCB by oxidizing both surfaces in oxygen plasma and bringing them into conformal contact.

Access holes were punched out of the elastomeric layers and aligned to fluidic ports in a PMMA plate. For increased structural stability, an additional PMMA plate was attached to the bottom. The multilayer structure was tightened with screws providing sufficient backing for the flexible PDMS layers.

An on-chip liquid reservoir was filled with small amounts of water and green dye under vacuum. As sufficient power is applied to the system, the composite locally expands upon dissipated heat from the integrated heater and liquid stored in the reservoir is consequently released into the microchannel as shown in Fig. 56.

Moreover, liquid flow in microchannels is entirely blocked by means of integrated valves. The devices are fabricated using low cost materials only and the concept allows for wafer-level processing. This microfluidic liquid handling concept with on-chip active components is highly integrable and does not require external actuators.
12. Electro-active polymer microactuators

Electroactive polymers (EAP) are a novel class of active materials that can be used in micro-actuator systems. Commonly, the EAPs are divided into two classes: electronic EAP and ionic EAP. Electronic EAP encompass

---

28 Edwin W.H. Jager, Daniel Carlsson
dielectric elastomers and electrostrictive polymers [472]. They are dry materials often characterized by their high operating voltages. Ionic EAP on the other hand depend on mobile ions to provide the main actuation mechanism. These materials are generally wet and they require relatively low driving potentials. Ionic EAP comprise ion polymer metal composites (IPMC) [473,474], conducting polymers (CP) [475], carbon nanotubes [476] and electroactive gels. For more details on EAP actuators and devices, we refer to ref. [477].

Most of the micro-actuators based on electroactive polymers use conducting polymers (CP) as the active material. These will be reviewed in this section together with some other EAP miniature devices. Pelrine et al. describe an electrostrictive linear microactuator [478]. Kwok and co-workers have presented sub-millimeter ionic polymer metal composite actuators [479]. These were made by laser ablation of Nafton sheets. The smallest actuators that they were able to fabricate were 300 μm wide and 3000 μm long with 20-μm thick fingers. By improving the IPMC processing and using micromachining techniques the same group was able to reduce the size of the IPMC actuators to 100 μm wide, 1200 μm long and 0.4-μm thick [480] (Fig. 57).

Conducting polymers have a number of properties that make them attractive as EAP microactuators. Some advantages are the low operating voltages and low power consumption, biocompatibility [481,482], ease of manufacturing and integration with common microfabrication processes [483].

12.1. Conjugated polymer actuators

CP microactuators are based on conducting or conjugated polymers that undergo a volume change upon oxidation or reduction of the material. The first devices based on conjugated polymers as the active material were demonstrated in the early 1990’s, mostly as bilayer devices [475,484–490]. One layer of the bilayer was typically a passive substrate onto which the electroactive conjugated polymer layer was applied, commonly by electrochemical deposition.

These bilayer devices were initially used to determine the amount of volume change and to identify the volume change mechanism [490,491]. Bilayers thereby provided a simple way to study conjugated polymers, such as polypyrrole, polyaniline [492,493], and polythiophenes [494]. Under electrochemical oxidation and reduction, the volume change of the active layer causes the entire assembly to bend. The direction of the volume change as well as its magnitude can be deduced from the direction and degree of bending. Such bilayer structures were typically a few centimeters long and a few millimeters wide. Studies showed that the volume change in the conjugated polymer is dominated by ionic movement into and out of the polymer. Fig. 58 illustrates a few of the possible actuation modes.

Other studies have been done on single conducting polymer fibers or strips submerged in a liquid electrolyte and connected to force/elongation measuring equipment. Speed of actuation, stress and strain were measured, verifying the results from the earlier bilayer studies [487,488,495,496].

As mentioned above, the volume change of conjugated polymers is controlled by the electrochemical processes that cause ion insertion and de-insertion. There are two mechanisms for the ionic flow. For a polymer P doped with large, immobile anions A+ in contact with an electrolyte containing small mobile cations M+

\[ P^+(A^-) + M^+(aq) + e^- \rightarrow P^0(A^-M^+) \]  
that is to say, cations can be inserted and de-inserted into the polymer. For a polymer P doped with small, mobile anions \((a^-)\) in contact with an electrolyte containing both mobile cations and anions:

\[ P^+(a^-) + e^- \rightarrow P^0 + a^-(aq) \]  

In this case, ions inserted and de-inserted are anions. In the former case, the volume expands in the reduced state of the polymer (a negative potential), and in the latter case the volume typically expands in the oxidized state (a positive potential). In the latter case, however, there may be two moving species. This is because not only reaction (2) but also reaction (1) may occur. This can lead to “twitching” behavior [491]. For a smooth actuation movement, it is preferred to have only one ionic species transferred. For microactuators, polypyrrole (PPy) doped with the large immobile anion dodecylbenzene sulfonate (DBS) has most commonly been used. This system provides smooth motion, as only cations are used as the moving species, as well as stability and a long life time.

The volume change is brought about by changing the applied electrical potential; in this way, changing the oxidation or reduction state of the conjugated polymer. This is not just a bistable system; however, intermediate potentials can also be applied, leading to intermediate redox states and this can produce intermediate volume changes.
and intermediate bending angles or strains. The potential can be applied both statically and dynamically and the 
volume change can be repeated a large number of times.

Multiple mechanisms are involved in the physical swelling of CPs. The macroscopic volume change of the polymer 
is mainly due to the combined effect of the insertion of ions and their solvent/hydration shell, which occupies free 
space within the polymer matrix [490,491]. Moreover, there is also an inflow of solvent, governed by osmosis, which 
acts to balance the altered ionic concentration inside the polymeric matrix [497]. Thirdly, there are conformational 
changes and coulombic repulsion acting on the existing polymer backbones which provide an additional contribution 
to the structural changes [498,499]. The degree of volume expansion and its speed is thus dependent on a number of 
factors. These include ion types and sizes, ionic concentration, the solvent, the thickness of the CP film and the external 
voltage applied [497,500–502].

A critical factor for the performance of CP actuators is the ion transport through the film which controls the swelling 
behaviour. The ion transport process is controlled by ion diffusion but also involves ion migration in response to the 
electric field [503].

In order to activate the actuators an ion source/sink is needed. Aqueous electrolytes have mainly been used for this 
purpose and hence these systems are often referred to as being wet. These fluids can be salt solutions such as those 
containing NaDBS or NaCl, and they have analogues in physiological solutions such as cell culture media and blood. 
In some cases it is preferred to have a “dry system”, that is a system that can be operated in air, instead of a liquid. 
Several ideas for this have been proposed and polymer electrolytes have been demonstrated in macroactuators 
[504,505]. In these devices, the polymer electrolyte is sandwiched between two CP layers, which generate the bending 
force. Encapsulation of a complete device using the CP layer, a hydrogel electrolyte, and a counter electrode has been 
demonstrated [506].

The first PPy based microactuators were PPy/gold bilayer devices made on silicon wafers using standard 
photolithography. These “fingers” with lateral dimensions in the range of a few millimeters curled when activated 
[507]. After this initial demonstration, the fabrication process was further developed and the PPy-microactuators were 
integrated into more complex microsystems such as self-assembling boxes [508], cell clinics [509] and microrobots 
[510]. A millimeter-scale actuator using polyaniline as the conductive polymer has been presented by Zhou et al. 
[480]. Using standard microfabrication techniques they fabricated bending actuators with dimensions of 0.5 mm by 
2.5 mm and 1 mm by 5 mm.

12.2. Fabrication of PPy-microactuators

Standard photolithography and microfabrication methods, including surface and bulk micromachining, are used to 
fabricate PPy-microactuators. For more details on microfabrication of conducting polymers we refer to reviews such 
as [511–513].

Conjugated polymers can be deposited in several ways such as casting, dip coating, or spin coating. However, in 
many cases this approach is limited by the insolubility of the polymer. Alternatively the polymer can be electro-
synthesized in-situ from a liquid electrolyte containing the monomer and a dopant salt. The electro-polymerization 
and deposition of conjugated polymers occurs via monomer oxidation, monomer addition or dimerization and 
subsequent growth of oligomers and polymers of the oxidized monomer [514]. The chain precipitates out of solution 
and forms a deposit on the electrode surface where the initial oxidation occurred. The polymer is deposited in the 
oxidized form, with compensating counter ions incorporated from the electrolyte.

Typically, for the PPy-microactuators, PPy(DBS) is electrochemically deposited onto Au-coated Si wafers from an 
aqueous electrolyte containing 0.1 M pyrrole and 0.1 M NaDBS. To initiate growth, a potential between 0.5 and 0.6 V vs. 
Ag/AgCl is applied to the electrode. A higher potential results in faster deposition but also in non-uniform thickness. 
As a counter electrode, another Au coated Si wafer is used, and as a reference electrode, an Ag/AgCl electrode.

The lateral dimensions of the conjugated polymer components are determined by the patterning step. Standard 
photolithography can be used in many ways to define the lateral dimensions (Fig. 59). One can pattern holes or 
openings in a layer of photoresist that is applied on a conducting surface to protect parts of the surface. The polymer 
only “grows” on the surface areas that are exposed to the solution. Alternatively one can synthesize the conducting 
polymer on the whole surface area and remove parts of the polymer by covering the layer with photoresist and 
subsequent Reactive Ion Etching (RIE) in an O2/CF4 plasma. Also one can pattern the electrodes areas first and 
polymerize the conducting polymer on these patterned electrodes.
The lower limits to the lateral dimensions of the microactuators are currently enforced by the photolithographic equipment. The smallest actuators that have been made were 10 μm wide and 40 μm long.

After fabrication of the actuators (or complete microsystems) the actuators have to be released. For the release of PPy microactuators three methods have been demonstrated: Differential adhesion, sacrificial layer, and bulk etching (Fig. 60). The differential adhesion method [515] is based on the poor adhesion between Au and Si. A Cr layer on a Si surface is patterned, resulting in adhesive and non-adhesive areas. Over this surface, an Au layer is deposited, and PPy is electrochemically deposited as described above. The PPy is patterned so that a minor part of the Au/PPy bilayer is in contact with the adhesive Cr layer, and this part functions as an anchor holding the actuator to the surface. The major part of the bilayer is in contact with the Si, to which the Au does not adhere. Activating the bilayer causes it to pull itself free.

Although the differential adhesion method is easy and fast, it is not suited for all designs. For example, it cannot be used to fabricate some individually controlled microactuators. To make individually controlled actuators, the metallic layer of the microactuators is patterned into segments, one for each actuator. These segments are mechanically connected by a resin layer (Benzocyclobutene, BCB, which is a rigid, transparent, conventional polymer similar to polyimide for micromachining applications). This step will cause the resin layer to be in contact with the Si-substrate, and as this resin is adhesive, the microactuators cannot release. For instance, to fabricate the micro-robot [510] the
sacrificial layer method had to be used. This is a common method to release object in microsystems. On a SiO₂ covered wafer, a support layer of Ti (the sacrificial layer) was patterned. Next, Cr, Au, and PPy layers were deposited. The microactuators were patterned with a small part overlapping the Si as the anchor point and the rest overlapping the sacrificial layer. Removing the sacrificial layer by under etching resulted in a free-hanging bilayer that could be activated. The advantage of the sacrificial layer method is that one can build more complex microactuator devices. Disadvantages include a long under etching time and potential damage to the PPy, depending on the etchant.

The third method is bulk etching. First, the PPy-microactuator or the complete microsystem is built. In one of the last steps a part of the substrate is removed by etching thus freeing the actuators [516,517].

12.3. Operation and performance

In order to actuate the devices an electrochemical cell has to be used. This comprises a control unit, an electrolyte, a working electrode, that is the PPy-microactuator(s), a counter electrode, and possibly a reference electrode. As a counter electrode typically an Au-coated Si wafer, Au wire or Ti/Pt mesh is used. As a reference electrode a commercial Ag/AgCl reference electrode can be used. For miniaturization all these electrodes can be integrated on-chip [518]. A microactuator with 100 µm by 150 µm Au counter electrode and a 50 µm by 100 µm Ag/AgCl quasi-reference electrode has been demonstrated. For the control unit one can use apparatus ranging from a simple battery to commercial research grade potentiostat/galvanostats. Also, integration of the potentiostat in the device that comprises the PPy-microactuators is being developed [519].

The performance of PPy microactuators has been characterized for microfabricated bilayer actuators. Bending angles, strain, speed and force are some important performance metrics [516,520]. For bilayer actuators the performance is largely influenced by geometric parameters. The length and width of the actuator are naturally influential as well as the total bilayer thickness and the individual thicknesses and ratios of the PPy and Au layers. For PPy(DBS) microactuators the optimum thickness ratio of PPy:Au for minimum curvature has been found to be 5:1 [521].

For good actuator performance and long lifetime, it is clearly necessary to avoid delamination of the PPy from the underlying substrate. Significant interfacial stresses are generated when the PPy expands and contracts [522]. Adding porous layers of Au has been a successful method promoting mechanical interlocking of the different layers [521,523].

The actuation strain of the PPy is a critical parameter that characterizes the performance of the actuators. For bilayer actuators the strain has been calculated based on mechanical models to be 0.45–3.4% in polypyrrole doped with various anions [490,492]. Bay et al. have also measured strain directly and reported strains of up to 4% for PPy doped with various alkyl benzene sulfonates [500]. The strain varies for different film thicknesses and a strain gradient through the PPy film thickness has also been reported [521].

Bending is a response only to in-plane volume change. The volume change of PPy has shown to be highly anisotropic. The out-of-plane volume change has been measured for PPy(DBS) and an expansion of >35% in the reduced state has been reported [524]. This out-of-plane volume change is sufficiently large that it can in principle be exploited directly in bulk type microactuators that do not need to be released from the surface. Such devices could be used as micropumps and valves.

Force measurements on PPy microactuators have been performed by measuring the mass that microactuators could reversibly lift. For a device with a PPy thickness of 0.9 µm, the ratio of the mass lifted to the mass of the hinge was 43,000:1. The radius of curvature was 5.6 µm, and it took 2 s to bend or straighten a bilayer 1 µm thick. Energy efficiencies were ~0.2%. [516].

12.4. Applications and devices

12.4.1. Bending actuators

The simplest devices are benders or fingers: bi- or multilayer strips of PPy/Au or PPy/Au/polyimide that curl and straighten under activation [507,525]. These can be used to grab small objects. For example, the cylindrical object in Fig. 61 could be a synthetic analogue of a nerve fiber.

In these fingers rigid beams can be added at different angles relative to the long side of the actuator in order to control the motion [526]. These stiff beams can turn the curling fingers into spirals with either a right hand or left hand turn (Fig. 62).
One can add other micromachined elements to the PPy-microactuators. This will result into more complex structures. Rigid plates can be added to the ends of the bilayers so that they function as hinges. These rigid plates can be made of BCB (Benzocyclobutene) [527], SU8 [509], or “bulk” silicon showing that one may potentially move and position anything that can be fabricated on silicon [516]. The plates can be rotated from 0 over 270° until they hit the surface of the substrate, depending on the device design, and might be used as mirrors or to change the surface of a device by exposing either one or another side of a flap. They have been combined in series, generating self-opening and closing boxes (Fig. 63) [527].

It is possible to deliver power to the plate and for instance build a moving pixel [528]. Lids can also be constructed to close and re-open microcavities or microvials that might contain drugs or cells. [509]. These sealable microcavities might also be used as small vessels for nano- and picoliter chemistry. The cell clinic is one such device. This is microcavity with integrated electrodes on the bottom and a lid operated by PPy-hinges (Fig. 64). It has been used to...
study cell behavior. Frog melanophores (Xenopus laevis melanophores) have been added to the cell clinic. It was possible to determine cell spreading on the surface and covering of the electrodes at the cell clinic bottom, and subsequent measurement of intra-cellular effects of stimuli added to the cells [529]. This work is continued with the addition more sensors to the cavity, resulting in what now is called a laboratory-on-a-chip [530].
Several individually-controlled actuators can be combined in a single device to create more complex structures. A microrobot has been fabricated and operated (Fig. 65) [510]. PPy microactuators were combined with stiff elements formed from the resin BCB to generate a “robot arm” consisting of an “elbow”, a “wrist”, and a “hand” with 2–4 “fingers”. The arm was only 670 μm long. With this robot arm 100 μm glass beads could be picked up, lifted, and moved over a surface. In another design, two connected hinges have been placed perpendicular to each other to create an actuator that could move in plane, perpendicular to the surface [531]. By intelligently combining such microactuators or hinge elements, new activation schemes in 3D space are possible.

A microgripper for underwater manipulation based on IPMC is under development by Zhou and co-workers, although these have not yet reached the small dimensions of PPy-microactuators [480].

12.4.2. Valves

PPy-microactuators are also commonly used as the active element in microfluidic valves. Several constructions have been demonstrated. The afore mentioned plates activated by hinge-type PPy-microactuators have been used to control flows in microchannels made of a silicone rubber (PDMS) slab assembled on a silicon wafer [532] and to open containers for a drug delivery device (“smart pill”) [533].

The out-of-plane volume change in these materials is found to be higher and this has been exploited in a valve [533,534]. In ref. [533], it was proposed as the valve for the “smart pill”. In ref. [534], it was used to open and close a microfluidic channel made in silicone rubber. The microfluidic channel was fabricated in assembled silicone slabs. The PPy-microactuator was a piston-like device mounted under the channel. Upon activation the PPy-piston pushed the bottom wall of the channel up, thus blocking the flow. An advantage was that the electrolyte chamber was separated from the flow channel. Yet, another type of PPy valve is used in an implantable drug delivery device called IntelliDrug (http://www.intellidrug.org) [535]. Here, a buckling PPy membrane is used to control a flow in a microfluidic channel, in order to control the drug release.

Yamada and co-workers have also demonstrated a PPy-microactuator based on the bulk expansion of PPy [536]. A 30-μm thick PPy film was synthesized on comb shape electrodes and covered with a solid polymer electrolyte. Upon activation the PPy expanded 2 μm.

The nature of CP microactuators makes them good candidates for applications in medical devices, in-vitro diagnostics and biomedicine [537]. Medical device applications, for example for use during vascular interventional procedures are for instance being developed by companies such as Micromuscle AB. The CP microactuators are being incorporated as parts of medical devices and enable new functionality by performing a range of functions, such as controlling the movement and mechanical properties of medical devices, holding and releasing objects inside vascular areas, etc. (Figs. 66 and 67).

The knowledge of CP material properties and of the mechanisms that control actuator performance increases constantly. This development will enable new and improved CP actuators for future (commercial) applications.

Fig. 66. EAP-steerable vascular device for maneuvering in vessels, demonstrated in a bench set-up.
13. Electrochromic and electroluminescent polymers

The discovery of conductivity in plastic materials has been a significant breakthrough in the field of engineering science since it has led to new avenues and applications. For example, a special type of electro-conductive plastics known as conjugated polymers are today emerging as a material for the production of electrochromic (EC) and electroluminescent (EL) devices. These devices are already at or near market for thin flexible displays, “smart” windows, rear-view mirrors, chameleon camouflaged clothes, light-emitting diodes (LED), field-effect transistors (FET), chemical sensors, solar cells and even coloured fingernails. The main reason for such a broad area of application is the versatility of polymer science and organic chemistry that enables a wide range of functionalities. From the market point of view, EC and EL devices are potentially much lower cost compared to common inorganic semiconducting devices, mainly due to easier processing technologies of the polymers. An expected price decrease per polymer EC/EL device in the near future and a large market potential attracted the attention of academia, large international companies and small start up companies.

Both phenomena, electrochromism and electroluminescence in polymers, are based on the unique property of the conductivity of conjugated polymers. Conjugated polymers have a backbone composed of alternating single and double carbon-carbon bonds. Single bonds known as $\sigma$-bond are made by overlapping two hybridized $sp^3$ orbitals, while double bonds known as $\pi$-bonds are made by overlapping two hybridized $sp^2$ orbitals. The remaining out-of-plane $pz$ orbitals also overlap each other, forming an electron cloud above and below the main backbone. It is depicted in Fig. 68 for the case of ethylene molecule. This cloud is thought to contain delocalized electrons capable of carrying charge, if polymer is exposed to the external field.

In comparison to solid state electronics, this situation is similar to the electronic structure of common inorganic semiconductors, for example silicon. Hence, we can now consider two energetic bands of conjugated polymers, the valence band ($\pi$-state), also known as high occupied molecular orbital (HOMO) and the conductive band ($\pi^*$-state) which corresponds to the lowest unoccupied molecular orbital (LUMO). The conductive band has a larger energy potential than the valence band so that electrons must jump from the lower to the upper band. The energy gap, $E_g$, is a fingerprint of every conjugated polymer that is dependent on the chemical composition and steric architecture of the polymer chain. Good knowledge of organic synthesis is the main key for fine-tuning of the $E_g$ and the resulting electrical properties. Typically, most conjugated polymers have the energy gap between 1.5 and 3 eV, which corresponds with visible to near-infrared electromagnetic radiation, i.e. from 400 to 1200 nm.

On application of a voltage to the polymer it can be switched between a neutral (undoped) state and a conducting (doped) state. Some conjugated polymers show distinct electrochromic behaviour in that the doped and undoped states have different colors. The mechanism of reversible colour switching in electrochromic materials is based on oxidative (loss of electrons, ’$p$-doping’) and reductive (gain of electrons, ’$n$-doping’) processes occurring at the anode/cathode interface. Upon electrochemical doping the band structure of the neutral polymer is modified, generating lower energy intraband transitions and creating charged carriers, which are responsible for increased conductivity and optical modulation.

---

29 Zoran Djinovic
Electroluminescence is an optoelectronic phenomenon whereby the material emits light when an electrical current passed through it, usually in response to a strong electric field. The main mechanism is radiative recombination of electrons and holes in material, e.g. as in a semiconductor [542]. When electrons and holes are injected from electrodes into an organic film or several organic layers, organic electroluminescence occurs and the devices are called organic light-emitting diodes. The cathode injects electrons in the lowest unoccupied molecular orbital (LUMO), i.e. in conduction or $\pi^*$ band of the polymer. Simultaneously, the anode injects holes in the highest occupied molecular orbital (HOMO), i.e. in valence or $\pi$ band of the polymer. Under the voltage these charges, named polarons, travel towards the opposite electrodes and are mutually annihilated by multi-phonon emission or surface recombination. As a result of the electron–hole interaction, a neutral space (exciton) has been formed. The exciton can be in the singlet or triplet state, depending on spin state. Only the singlet state can contribute to the photon emission and since the singlet state to triplet state ratio is 1/3, the maximum quantum efficiency with luminescent polymers can be 25% [547]. The remainder of the electrons will annihilate by some non-radiative means, for example in causing photo-oxidation of the polymer.

13.1. Electrochromic materials

Early electrochromic (EC) materials and devices were demonstrated on inorganic materials such as metal oxides and wide-band-gap semiconductors, mainly WO$_3$ [548,549]. Electrochromism in organic materials was initially discovered using small molecules such as bipyridiliums (viologens), but it has also been observed in conjugated polymers such as those compounds based on poly(thiophene) (PTh), poly(pyrrole) (PPy) and poly(aniline) (PANI) [550]. These materials are the prime candidates for display technology, due to their high contrast ratio (electrochromic contrast), short response times (switching speed) and long lifetimes (stability) [541].

There are three main groups of EC materials. The first group comprises materials with two distinctive optical states, colored and bleached (transmittive) state. They find their application in so-called absorption/transmission EC devices such as “smart” windows and optical shutters. The most frequently used organic polymer in this group is poly(3,4-ethylenedioxythiophene) (PEDOT). The second group encompasses those materials used for display applications, since they have two clearly defined colored states. For example, PTh can change colour from red to blue under oxidation [541]. The third group consists of materials known as multi-colour polymers. These have more than two optical states such as poly(3,4-propilendioxypyrrole) (PProDOP) and the above-mentioned PANI. These materials also find application in display technology as conductive layers [551].

One of the main advantages of the conjugated polymers, particularly multicolour polymers, is the unlimited possibility to tailor EC properties by changing the polymer structure. A large number of synthetic techniques are involved for fine-tuning of the $E_g$ of conjugated polymers [552] and ultimately their response to the applied electric fields. In the simplest method, it is possible to substitute the parent heterocycle causing steric and electronic influence on the $E_g$ of polymer. Homo-polymerization of co-monomers or co-polymerization of different monomers also results in modification of the main backbone of the polymer, providing it with new properties. Electro-synthesis has been frequently used to produce structurally different polymers [553]. In addition, to these chemical techniques there are several physically based techniques for altering the electrochromic properties of the polymer, for example, polymers can be utilized in blends, laminates or composites [554–556].

A large number of different compounds are produced by relatively small structural changes of the starting monomers such as PTh, PPy and PANI, providing new colours in their neutral or doped forms. For example, Gaupp et al. [557] slightly modified the structure of the PPy to include an additional aromatic ring and they obtained several new derivates, such as PEDOP and PProDOP with $E_g$ of 2.0 and 2.2 eV, respectively. Hence, PEDOP gained in its
neutral state a red colour, while in its oxidized state the polymer was transmissive blue. Moreover, PProDOP achieved three coloured states, an orange neutral, an intermediate brown and a grey/blue oxidized state. The effect of substitution of the H atom in the NH-group with a sulfonato-propoxy group was even stronger causing a significant increase of the $E_g$, to more than 3.0 eV in N-PrS-PProDOP (poly(N-sulfonatopropoxy-ProDOP)) derivate. Consequently, the polymer changed its colour from its transmissive colourless neutral state to an absorbing light grey oxidized state [558]. This was a direct consequence of the steric interaction of the rather bulky substituted groups with the rest of the polymer chain. The involvement of more rings into the basic repeated unit, as well as combining the thiophene and pyrole rings, drastically increased the number of possible derivates. Depending on the doping level, they demonstrated multi-colour ability, resulting from variability in the energy gap, from 1.2 eV with PBEDOT-PyrPyr (poly-bis-EDOT-pyridopyrazine) to more than 3 eV [559]. By inclusion of the Si atom as a link between the two bithiophene rings in the repeated unit, an extremely low $E_g$ of about 0.85 eV was obtained [560] enabling the material to be used for near infra-red (NIR) applications. This significant decreased of $E_g$ is explained by shortening of the conjugated length of the polymer chain.

13.2. Electrochromic devices

Electrochromic devices (ECDs) have the sandwich structure that is schematically presented in Fig. 69a. Fig. 69b shows the electrochromic effect of one absorption/transmission-based ECD composed of PProDOT-Me2 and N-PrS-PProDOP polymers [541]. The active electrochromic material is coated onto a transparent glass or plastic substrate and driven by a low voltage source of ±1.5 V. An electrochemical reaction (redox reaction) then takes place between the cathode and anode via oxidative or reductive processes. Electrochromic behaviour depends on the diffusion of ions from one layer to another.

ECDs based on inorganic materials are relatively slow to change to their two-colour states, with timescales that are typically of the order of several seconds. Liquid crystal displays (LCDs) are much faster, but they suffer from a limited viewing angle and they lack the multi-colour effect. Hence, the current need in the display technology is to have much faster, stable and higher contrast displays. Based on the latest developments, electrochromic polymers are promising materials that can meet the above mentioned demands. For example, Aubert et al. [561] were able to achieve a switching rate in the order of 5–10 Hz, with about 60% of contrast ratio in the visible region and they maintained stable ECD performance over 180,000 cycles. The polymers used included those mentioned above, such as PXDOT poly(3,4-alkylenedioxythiophene), PEDOT, PProDOT as well as some derivates of these polymers, changing blue and red neutral state colours to a bleached oxidative state.

Despite the evident successes, the problem of long-term stability has remained a major challenge for the polymer based ECDs. However, the use and development of novel classes of polymers, blends, copolymers, composites and laminates and new ionic media have enabled fabrication of ECDs with increased stability against environmental conditions especially in air and oxygen interaction. Furthermore, the colouration efficiency, i.e. the amount of charge required to produce the optical change, will in all likelihood reach 500–3000 cm$^2$/C very soon [541]. Finally, it is possible to contemplate the production of large-area flexible ECDs due to relatively simple processing technology involved. This is primarily based on solutions of the polymers and readily compatible with current ink jet technology.

Fig. 69. (a) Schematic presentation of an ECD and (b) example transmittance spectra of absorption/transmission electrochromic devices [541].
In practice there are two main types of ECDs, the absorption/transmission (A/T) type and the reflective type. The first group is composed of two complementary polymers, deposited onto the cathode and anode, respectively, and separated by an electrolyte. Usually, this is a viscous gel. The electrodes are typically made of a transparent solid such as indium tin oxide (ITO)/glass or flexible ITO/polyethylene terephthalate (PET) materials. When a voltage is applied between the electrodes, a high-band gap polymer becomes oxidized by the anode and changes the transmissive state to a coloured state. At the same time, a low-band gap polymer, which is deposited on the cathode and is coloured in the neutral state, becomes transmissive upon oxidation. Hence, the ECD switches between two states, coloured and transmissive. This is the main functionality necessary for “smart” windows and optical shutters. Fig. 69b shows these events when two complementary polymers PProDOT-Me₂ and N-PrS-PProDOP have been driven by only ±1.5 VDC [541]. Reflective ECDs have a similar structure to that presented in Fig. 69a with the exception that one electrode is non-transmissive, usually made of Mylar foil coated by gold. When a thin film of PProDOT-Me₂ is applied on the gold electrode it is possible to obtain switching from a deep blue to a reflective gold state. However, these devices can be active not only in the visible region but also in the near-infrared, mid-infrared and even micrometer regions, yielding a contrast ratio as high as 90% at 1.8 μm. This functionality is very important for thermal emissive control, often demanded in military applications, for example for camouflage purposes [562].

High volume production of low-price ECDs will be feasible in the near future due to comparatively simple polymer production technologies. For example, the solvent state of the active polymers is compatible with well known and mature screen printing techniques with resolution values of about 10–100 μm. Micro-contact printing (μCP) and inkjet printing can also be utilised [563]. Fine electrode structures can be obtained by line patterning, which combines metal-vapor deposition and inkjet printing [564]. ECDs most certainly have a large market potential, not only in microtechnology, but also for much larger scale applications. An example of this would be in the building industry for the production of so-called ‘smart windows’. Electrochromic windows can help save energy by automatically controlling of the light and solar energy that can pass through the window. The US Solar Energy Industries Association estimates that smart windows could save as much as 50% of building’s energy. This is based on the premise that they can lower cooling loads in the summer months by reducing solar heat gain [565]. At present, electrochromic windows are usually made of inorganic EC materials [566].

13.3. Electroluminescent materials

Electroluminescence (EL) is the generation of light in response to an externally applied voltage. EL was first observed in an inorganic material, zinc sulphide, in 1936 by Destriau [567]. In 1950, Bernanose observed the effect in organic compounds when thin films of acridine orange and qinacrine [568] were exposed to a high alternating voltage. The same effect was observed in 1963, when Pope et al. applied a voltage of ~400 V to an anthracene film which generated a blue light [569]. Leading on from these valuable initial observations, two-layer structures composed of p-type aromatic diamine and n-type of emitting layer of aluminium chelate tris-hydroxyquinoline (Alq₃), were created and this represented the first real example of organic light emitting diodes (OLEDs) capable of working at a relatively low voltage of ~10 V.

In 1990, Friend and co-workers [570] made a dramatic breakthrough in the field of EL materials reporting that the conjugated polymer poly(p-phenylene vinylene) PPV showed green-yellow electroluminescence with \( E_g \) of ~2.5 eV. Although, the first simple device had an EL efficiency (the number of generated photons per number of injected electrons) of only 0.01%, this discovery heralded the beginning of polymer light emitting diode (PLEDs) or OLED technology. Braun and Heeger also discovered that a PPV soluble derivate, poly[2-methoxy-5-(2‘-ethylhexyloxy)-1,4-phenylenevinylene], or MEH-PPV in short [571, 572] generated a red EL colour with an \( E_g \) of ~2.2 eV. Since MEH-PPV was a soluble conjugated polymer it offered a number of advantages in processing technology. However, it rapidly became clear that this particular EL polymer is highly sensitive to photo-oxidation [573], which was a significant disadvantage for further development.

PPV was soon recognised as being unsuitable for the fabrication of PLEDs, mainly due to difficulties of processing the solid phase of the polymer. Significant effort has been made within the scientific community [574–576] to synthesise a number of soluble PPV derivates involving different side groups such as alkoxy, alkyl and cyano groups. It was also shown [577] that side groups have a significant influence on the magnitude of the energetic gap, electron and hole mobility and luminescence efficiency. As a result, the PPV derivate, MEH-PPV, gained about 1% of the EL efficiency at 3–4 V, due to alkoxy substitution in side chains of the PPV backbone. Alkyl substitution produced a shift
of the luminescence colour from yellow to blue with an increase of efficiency of approximately 2.5% [578]. Cyano side groups improved electron mobility in CN-PPV derivatives, while triphenylamine and carbazole molecular fragments increased hole mobility.

Poly(p-phenylene) PPP is a new EL active polymer, derived from PPV, which has a higher $E_g$ between 2.8 and 3.5 eV. PPP materials have received a lot of attention with regard to the creation of blue PLEDs. Unsubstituted PPP, like PPV, is insoluble in organic solvents and this makes fabrication procedures for PLEDs difficult. Trials [579] to deposit thin PPP film by vacuum deposition have been undertaken, but generally this technique requires a pre-polymeric precursor. Like PPV, PPP can either be modified in soluble derivatives by incorporating of alkyl-, aryl-, alkoxy- or perfluoroalkyl-groups. Sulfonate-, carboxy functional groups and ammonium have the effect of making PPP water-soluble [580].

In the past 10–15 years, there have been over a thousand different polymers [581] developed for ELD applications, most of these are based on parent monomers such as thyophene and phenylene. The goal has been to tailor the colour of the EL polymer with similar methods to those described above. For example, Yoshio and co-workers prepared the wide-bandgap polymer polyfluorene to emit a rich blue EL colour [582]. Researchers at Dow Chemical Co. [583] synthesized a family of 9,9-diarylfluorenes with a molecular weight in the range of 50–600 kDa, indicating a degree of branching in the polymer chain, with $E_g$ of 3.68 eV and a blue EL colour corresponding to a primary emission peak at 436 nm. Many high-efficiency, high-brightness and long-living devices have been constructed from these materials. For example, Friend et al. [584] succeeded in fabricating an optimized green LED with very pure fluorine-based polymers. They obtained light emission of 10000 cd/m$^2$ by applying less than 6 V. Peak efficiency was 22 lm/W. Table 9 presents the most important achievements obtained for three different colours at Cambridge Technology Display.

At the same time, collaboration between Donetsk University, Ukraine and the University of Durham in the UK, has produced new stable variants of PF. The Durham–Donetsk collaboration replaced the carbon at the 9-position in PF with SO$_2$ to prepare dibenzothiophene-S,S-dioxide, which was then formed into co-oligomers with 9,9-diethoxyfluorene [586]. This improved electron transport and stopped it degrading. Kim et al. [587] recently achieved high purity of the blue color emitted by the poly[2,7-(9,9-di-n-dioctylfluorene) (PF) by incorporating gold nanoparticles (Au NPs) into the host PF polymer. At the same time, the stability of the blue color was greatly improved by Au NPs that suppressed the photo-oxidative and degenerative processes of the PF.

Recently, Shen et al. [588] reported the combination of two polymers, PF and MEH-PPV in order to get efficient white-polymeric light emitting diodes (WPLED) with brightness from 200 to 10,000 cd/m$^2$ as the forward bias increases from 5 to 10 V. The EL turn-on voltage is below 4 V and the maximum luminance reaches 10,270 cd/m$^2$ at the bias of 10 V. The luminous power efficiency and luminous efficiency were found to be about 1.8 lm/W and 2.3 cd/A, respectively for the bias of 4 V.

### 13.4. Electroluminescent devices

Basic device architecture is presented in Fig. 70a. This shows a polymer thin film, between two electrodes on a substrate. Fig. 70b presents the electroluminescent spectra of previously mentioned MEH-PPV in combination with PMMA. An OLED or PLED consists of a light-emitting polymer film, an optically transparent anode and a metallic cathode, together with a dc or ac power source. The anode is most often indium tin oxide (ITO) coated glass. This material has a relatively high work function (4.8–5.2 eV), and is a good material for injecting holes into the active polymer layer. The hole mobility in PPV is of the order of $10^{-4}$ cm$^2$ V$^{-1}$ s$^{-1}$ allowing transport across a 1000 Å layer with typical fields of 5 × 10$^5$ V cm$^{-1}$ in 0.5 μs. The cathode material is typically a low work-function metal, such as Ca, Mg or Al. For example, Ca has about 2.8 eV of work function that is very close to the value of LUMO of PPV. They are very efficient at injecting electrons from the cathode into the polymer layer. The polymer may be deposited on the ITO by spin- or dip-coating from solution, and is typically of the order of 100-nm thick. The cathode metal is evaporated onto the polymer film in vacuum. In order to improve injecting of holes there is very often an intermediate

#### Table 9

Performance for red, green and blue emitting polymers [585]

<table>
<thead>
<tr>
<th>Colour</th>
<th>Efficiency at 100 cd/m$^2$ (monthly average)</th>
<th>Voltage at 100 cd/m$^2$ (monthly average)</th>
<th>Lifetime initial from 100 cd/m$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>1.7</td>
<td>2.4</td>
<td>&gt;40,000</td>
</tr>
<tr>
<td>Green</td>
<td>9.0</td>
<td>2.7</td>
<td>&gt;20,000</td>
</tr>
<tr>
<td>Blue</td>
<td>2.0</td>
<td>3.9</td>
<td>3,500</td>
</tr>
</tbody>
</table>
layer between the ITO and active polymer, e.g. PEDOT polymer. The efficiency of electron injections from the cathode side could be achieved by deposition of one intermediate layer between the cathode and polymer, such as PF-N2 (aminoalkyl substituted polyfluorene). More complex structures of PLED can involve multilayer composition of several active polymers and one or more transport layers.

PLEDs are currently approved devices for high-volume production of low cost video and graphic displays. They can be processed in a relatively simple way by applying some of the well matured technologies such as spin coating or inkjet printing (see Fig. 71a) [584] onto the ultra thin, large and flexible plastic substrates. Using these techniques it is possible to define complex light emission patterns with high pixel resolution. Fig. 71b shows such a structure [590]. Completed devices provide 180° viewing angle without loss of aperture, backlight, colour filter and polarizer. The devices consume low amounts of power, working on dc current under the voltage lower than 5 V. Hence, they do not

Fig. 70. (a) Basic structure of PLED and (b) electroluminescent spectra of MEH-PPV with PMMA (Reproduced with permission - Professor John R. Reynolds, Center for Macromolecular Science and Engineering, University of Florida [589]).

Fig. 71. (a) Ink-jet printing of EL polymers [584] and (b) Integrated red-green-blue devices (Courtesy of Professor James C. Sturm, Princeton University – reproduced with permission © IEEE [590]).
generate any significant heat during operation. These properties, combined with a light, flexible structure and high switching speed, make the PLEDs highly attractive structures for innovative designs and new products.

Despite the promising results that have been obtained in reliability, efficiency and usability of EL polymers [567,576], there are still a lot of challenges that have to be overcome. The most important of these relate to: short lifetime, colour stability, modest efficiency and manufacturing technology. It is a well known that some of these organic compounds are very sensitive to their environmental conditions, for example, to moisture, air, heat or UV radiation [591]. The influence of these factors induces ageing of the polymer and deterioration of their useful properties, shortening the lifetime of the ELDs. The lifetime of an ELD is defined as time it takes for the device luminance to drop to half of its initial value. Typically, it has to be more than 10,000 h for most consumer electronic displays. Cambridge Display Technology Ltd. (CDT) has recently announced the achievement of an extrapolated lifetime of over 100,000 h for laboratory devices with brightness 100 cd/m² at room temperature [584]. These are colour displays composed of three adjacent pixels emitting red, green and blue light (RGB). Current red and green polymers can meet the stability requirements for displays. However, a significant challenge still exists in terms of the stability of the blue EL polymers. Finally, there are some particular problems concerning the fabrication procedure, such as patterning of the light emitting area, encapsulation of PLED and in the use of organic solvents [592].

Progress has been made in micro-fabrication of PLED pixel arrays by reduction of the active area to 0.8 μm, and by employing ultra-thin and self-assembling polymer coating techniques such as Langmuir-Blodgett deposition [593]. In this way, it has been possible to fabricate perfectly assembled layer-by-layer thin film multilayer hetero-structures. The main benefit has been precise control of the structural features and thickness of the single layers at the molecular scale. It is well known that on this scale, at the interface of the multilayer contacts, it is possible for quantum effects to emerge [547]. For example, a multi-layer structure composed of an aluminium cathode, sulfonium PPV and sulfonated polyaniline (SPAN) deposited on ITO (anode)/glass substrate emitted blue light instead of the greenish-yellow light that is typical for a single SPAn/PPV layer structure. This has been explained as being due to a confinement effect of carriers in the superlattice multilayer structure of SPAn/PPV [594].

13.5. Conclusions

Electrochromic (EC) and electroluminescent (EL) devices have already found applications for ‘smart’ windows and video displays. This has been made possible due to the unusual semiconducting properties of the conjugated polymers. Unlike inorganic semiconductors, these polymeric materials can be processed using simple and low cost techniques, such as spin coating, ink-jet printing and casting. This may result in a significant decrease of the production cost, increased flexibility in design and the ability to create devices in a large range of sizes. Furthermore, their low power consumption operating potentials below 5VDC, high switching speed, good colour brightness, a broad viewing angle of 180° and long lifetime in the range of 100,000 h (PLEDs) will enable these materials to compete for high volume production of graphic and video displays. Fine tuning of the electronic band gap and colour of the active polymer has been reached by structural modification of monomers, changing the composition of copolymers, involving substitutions and additions to the host polymer, e.g. Au nano particles, or simply by mixing two or more polymers. To-date, good purity and long time stability of the red and green colour of EL polymers has been obtained. However, some challenges remain to achieve stable blue-emitting colours from electroluminescent polymers.

14. Ferroelectrets—cellular piezoelectric polymers

Ferroelectrets are a new class of piezoelectric polymers. They consist of non-polar space-charge electrets with cellular foam structures. An electret is a material which can exhibit an internal polarization due to: (i) trapped charges, e.g. in non-polar space-charge electrets, (ii) oriented molecular dipoles in polar electrets, or a combination of both of these effects [595]. A well-known polar piezoelectric polymer is polyvinylidene fluoride (PVDF) in which the piezoelectric activity arises from the change of dipole density during the application of a mechanical stress or an electric field. Details of classical ferroelectric polymers are given in a book edited by Nalwa [596].

In the case of the new non-polar ferroelectrets, the microscopic origin of the piezoelectric activity is completely different compared to that of well known polar ferroelectric polymers. In ferroelectrets charge separation exists,
effectively creating upper and lower void surfaces that are oppositely charged (Fig. 72). This is the origin of the breakdown of symmetry, which is required for piezoelectric characteristics to emerge. Each void can be considered as a macroscopic dipole. An applied mechanical stress leads to the decrease of void sizes accompanied by the generation of an electrical signal. Therefore, piezoelectricity in ferroelectrets results from changes of dipole sizes.

The development of ferroelectret materials (previously referred to as cellular electro-mechanical films) was initiated in Finland during the late eighties [597,598] and strengthened in the late nineties [599,600]. Several groups entered this research field and today ferroelectrets are established as a major form of piezoelectric polymer, along with ferroelectric polymers. The mechanisms, preparations, properties and applications of ferroelectrets have been described in recent reviews [601–603].

To-date piezoelectric foams of polypropylene PP [602,603], polyethylene terephthalate PETP [604–606], a variety of compounds of cyclo-olefine polymers COP, copolymers COC [607–610] and Teflon® AF [611] have been developed. Ferroelectrets prepared from PP, PETP and COP form the basis of the following brief overview, which is focused on the preparation of cellular foams, the internal charging of the voids, the resulting piezoelectric properties and concluding with a discussion of demonstrated applications.

14.1. Foam preparation and optimization

In the design of a ferroelectret, a cellular void structure is required as the starting point to obtain piezoelectric properties from non-polar space-charge electrets. The optimal cellular structures consist of lens-like voids with diameters of approximately 10–100 μm and widths of approximately 1–100 μm. Typical cellular structures of PP, PETP and COP are shown in the scanning electron microscopy (SEM) images in Fig. 73, adapted from the references [612,604,609]. Recently, a non-destructive 3D-scan of a cellular ferroelectret was performed with optical coherence tomography (OCT), which yields additional information, e.g. about the form and dimensions of the cell walls [613].

Different foaming procedures with separate preparation steps were developed in order to obtain optimized cellular structures. The first step is the initial foaming obtained either by extrusion of filler loaded polymers or by treatment in gases or supercritical fluids (SCF). The extrusion method often leads to cracks and void formation during the manufacturing process (schematically shown in ref. [614]). The latter treatment utilizing gases, e.g. in carbon dioxide or supercritical carbon dioxide, during or after the extrusion leads to diffusion of the gas into the polymer. When the pressure is suddenly released, the gas expands violently, which leads to a voided structure. This process is supported by a heat treatment. Ferroelectret foaming by SCF and heat treatment is described in refs. [605,606] the principle of their methods and their application to other materials is described in ref. [615]. With this approach, approximately round
voids are obtained depending on the base material. After forming the porous polymer, lens-shaped voids are formed by a biaxial stretching of the foamed polymer films. The void heights (which represent the size of the macroscopic dipoles) can be subsequently adjusted by an inflation process [603]. A variety of inflation procedures have been studied in detail [616–622], where suitable inflation parameters, the resulting void structures and their influence upon the piezoelectric properties were examined. As a result, the void sizes in filled-loaded, extruded, stretched and inflated polymers can be adjusted and tailored from flat lens-shaped up to large round voids with diameters of approximately 100 μm [620].

The variety of cellular structures that can be formed result in different piezoelectric and elastic properties, which in turn depend significantly upon the foam porosity. The foam porosity can be determined by the film density. The interdependencies are shown schematically (Fig. 74a) along with experimental data in Fig. 74b. A low elastic stiffness of the foam is essential for a high piezoelectric activity (see Fig. 74a). For PP ferroelectrets the optimized void length and height ratio is approximately 4–5, which corresponds to lens-like voids. However, the cell-wall thickness must be taken into consideration [623]. Furthermore, there is a lower limit on the elastic modulus since deformation during use of the material as a transducer should be reversible. The above-mentioned experimental conclusions regarding manufacture of cellular materials are also supported by numerical calculations [624] and by recently performed experiments [625]. Therefore, these dependencies find great use in the development and application of new ferroelectrets.

14.2. Void charging in cellular space–charge electrets

Non-voided electrets were traditionally charged electrically by means of a corona discharge. Initially, the corona discharge method was also preferred for charging of foamed electrets and it is now the accepted method for charging of large-scale ferroelectret transducers. However, the detailed understanding of the mechanism of charging cellular structures was obtained by examining a range of charging techniques, along with the variation of charging parameters.
Cellular films are charged by corona discharge or direct application of high electric fields in the range from tens to hundreds MV/m. The large field across the film thickness leads to dielectric barrier micro-plasma discharges inside the voids, which results in trapping of charges with opposite polarities on the top and bottom sides of the voids, as schematically shown in Fig. 72. Evidence for the nature of this mechanism is provided by several experimental results. An electric field above a threshold is necessary to initiate the micro-plasma discharges, which were seen during charging experiments with varying electric fields applied by means of a corona discharge [626] as well as during direct-contact charging [614,627]. The electric charging depends significantly on the ambient atmosphere in the sample chamber and in the sample itself [605,612,616]. The micro-plasma discharges can be visually observed, and were recorded in situ during direct-contact charging via transparent electrodes [628].

Charge patterns were measured by scanning electron microscopy (SEM) on obliquely cut cellular PP films and this constitutes direct proof of the internal void charging [629]. Furthermore, the switching of the internal polarization was demonstrated by means of direct-contact charging [614,628], which gives rise to a hysteresis curve of the polarization or the piezoelectric activity as a function of the applied electric field. In summary, the charging of ferroelectrets is significantly different from the charging of non-cellular space-charge electrets. Charging of ferroelectrets leads to two different stable polarization states and is therefore comparable to the charging of ferroic materials, which was also the reason for the development of the name ‘ferroelectrets’ (for further details see the reviews [601–603]).

14.3. Piezoelectric properties

In general, ferroelectrets show high piezoelectric activity after suitable foaming and optimized electric charging. Typically, piezoelectric $d_{33}$ coefficients of several hundred pC/N or pm/V were reported for the direct and inverse piezoelectric effect in cellular PP films [616,619–622,630–632]. The recently developed PETP ferroelectrets and COC and COP ferroelectrets, which were prepared by extrusion of filler loaded polymer, exhibit a lower piezoelectric activity, at least under the current manufacturing conditions [604,605,607–610].

Ferroelectrets exhibit significant thickness-extension resonances [604,619,630,631]. Bauer and co-workers found length-extension resonances, but with a piezoelectric activity two orders of magnitude lower than for the thickness extension [631]. Furthermore, they showed that the pyroelectricity of ferroelectrets is very small and approximately two orders of magnitude below that of classical pyroelectric materials such as ferroelectric PVDF. Table 10 summarises typical values of the piezoelectric coefficient $d_{33}$, the elastic stiffness $c_{33}$ and the thickness-extension resonance frequency $f_{TE}$ for different ferroelectrets. An extended comparison of piezoelectric and elastic properties of PP ferroelectrets and ferroelectric PVDF is given in ref. [633]. A direct comparison of the $d_{33}$ coefficient of ferroelectrets and ferroelectrics such as PVDF can give a misleading view of their relative performance due to the large difference in their elastic stiffness, by a factor of 1000 or more (and therefore the dependence of their piezoelectric activity on pre-pressure), pyroelectricity and $d_{31}$ activity.

It should be mentioned that specialised preparations or treatments can lead to significantly higher piezoelectric coefficients of 790 [616] or 2000 pC/N [621] in PP and approximately 476 pC/N in PETP [606] ferroelectrets.

The thickness-extension resonance in the high kHz up to the low MHz region is the basis for the application of especially PP ferroelectrets in the audible and ultrasonic frequency range as microphones [632,634,635] or as loudspeakers [636–638]. A further advantage of PP ferroelectrets driven in the audible and ultrasonic frequency range is their low quadratic and cubic distortion, which is found to be typically 20 dB below the fundamental sound [634,636].

PP ferroelectrets show high piezoelectric activity, however, in the optimization of the transducer properties various approaches such as stacking of individual layers were attempted. A layer system can be arranged by folding or bending

<table>
<thead>
<tr>
<th>Ferroelectret material</th>
<th>Elastic stiffness $c_{33}$ (MPa)</th>
<th>Piezoelectric coefficient $d_{33}$ (pC/N)</th>
<th>Resonant Frequency $f_{TE}$ (kHz)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>1.3, ... , 14</td>
<td>up to 340</td>
<td>600, ... , 2000</td>
<td>[619]</td>
</tr>
<tr>
<td>PETP</td>
<td>7</td>
<td>up to 23</td>
<td>560</td>
<td>[604]</td>
</tr>
<tr>
<td>COC/COP</td>
<td></td>
<td>up to 25</td>
<td></td>
<td>[607]</td>
</tr>
</tbody>
</table>

Table 10
Piezoelectric and elastic properties of ferroelectrets
a ferroelectret film [639,640], by stacking or gluing ferroelectret films on top of each other [632,638] or connecting the ferroelectrets via a spacer air gap [608]. For PP and COP ferroelectrets, the piezoelectric activity increases linearly with the number of active layers [632,638], however, only for the first approximately 5–10 layers, since the foam weight eventually influences the piezoelectric activity of the lower films in the stack [609,638].

A critical property in terms of applications of ferroelectrets is the thermal stability of their piezoelectric activity. For PP ferroelectrets, the thermal stability of the high piezoelectric activity is limited to approximately 60 °C [631,641,642]. Annealing at elevated temperatures decreases the piezoelectric activity, due to loss of polarization within the voids and changes of void geometry. However, the resulting lower activity is thermally stable up to approximately 80 °C. Therefore, PP ferroelectrets can be made to fulfill some application requirements, as discussed below. In comparison to PP ferroelectrets, a slightly increased piezoelectric activity is found in PETP ferroelectrets [604]. A significant improvement in thermal stability is obtained with the cyclo-olefine polymers and copolymers. Excellent thermal stability up to at least 110 °C was demonstrated for some COP compounds [607,609].

For PP ferroelectrets in particular other environmental influences such as UV radiation and higher humidity were tested. UV radiation decreases the piezoelectric activity slightly [643]. Storage in higher humidity and at temperatures of up to 50 °C for 12 days leads to a 25% decrease of the piezoelectric coefficient [642].

14.4. Applications of ferroelectrets

Several applications were suggested and demonstrated for PP ferroelectrets as described in the reviews [601] and [603]. These applications are often similar to those of other thin piezoelectric films such as, e.g. ferroelectric polymers. However, the difference in the structural properties of ferroelectrets and ferroelectrics allows them to support different applications. A perceived disadvantage of ferroelectrets is the strong dependence that their piezoelectric activity has on the applied static pre-stress [637,642,644,645], whereas their very low pyroelectric activity is, in contrast, an advantage, being on a similar level to that of the ferroelectric PVDF. Thus, in ferroelectrets the piezoelectric signal will not be super-imposed with a pyroelectric current if a temperature change occurs.

Some examples of applications are quasistatic sensors for motion detection used in factory buildings for safety reasons, as well as for traffic control of cars and other vehicles as suggested in ref. [646]. Another application area is pressure monitoring for sports and orthopedic diagnostics [647]. Applications, which are already implemented and currently served by PP ferroelectrets, are push buttons for keyboards for a variety of equipment (for details see the description in ref. [603]). A new development in this area is the combination of ferroelectrets with integrated field-effect transistors [648].

As discussed above, PP ferroelectrets are suitable as microphones, which were already realized in the early nineties and used to record the respiration of human patients [649]. Furthermore, sensors working in the frequency range from 50 Hz to 23 kHz are produced to function as pickups for musical instruments by the company EMFIT in Finland. This application is probably the largest currently, in terms of manufactured quantities.

The application of ferroelectrets as sensors and actuators in ultrasonic applications such as material testing and industrial gas-flow measurements is a current area of research and development. An unusual ultrasonic application was realized with the implementation of PP ferroelectrets as a transducer to mimic the echo-detection capabilities of a bat in flight [635,650].

14.5. Conclusions and outlook

Ferroelectrets are a new class of piezoelectric polymers. They exhibit outstanding piezoelectric activities compared to ferroelectric polymers and enable applications such as quasistatic sensors and as transducers at audible and ultrasonic frequencies. Some applications such as keyboards and pick-ups for musical instruments are already implemented industrially. Until now, ferroelectrets of three different electret polymers, i.e. PP, PETP and compounds of different cyclo-olefine polymers or copolymers were developed. Extensive knowledge about the processing of PP ferroelectrets allows for a wide range adjustment of piezoelectric and elastic properties.

The new processing possibilities of PETP ferroelectrets by treatment with supercritical fluids offer better lab-scale investigation of the basic foaming mechanisms. The most thermally stable ferroelectret transducers are prepared from compounds of cyclo-olefine polymers. Investigations of suitable compounding of cyclo-olefine polymers are still
running. Furthermore, the expansion of the ferroelectret concept to fluorinated polymers is under development. Based on the excellent electret properties of fluorinated polymers [595], a high thermal stability of fluorinated ferroelectrets is expected.

15. Conductive polymers

Conductive polymers are sometimes called ‘synthetic metals’ because they present electric, electronic, magnetic and optical properties inherent to metals or semiconductors, while preserving the mechanical properties of conventional polymers.

Until about 30 years ago, all carbon-based polymers were rigidly regarded as insulators. The notion that polymers could be made to conduct electricity would have been considered to be absurd. They were used, for example, as inactive packaging and insulating materials. Alan J. Heeger, Alan G. MacDiarmid and Hideki Shirakawa changed this conception with their discovery that a polymer, polyacetylene, can be made conductive almost like a metal.

They managed to make polyacetylene films that were 109 times more conductive than they were in their original state by oxidation with chlorine, bromine or iodine vapour, and they received the Nobel Prize in chemistry for this achievement in 1977 [651].

Since that time some interesting applications of these materials have been developed [652,653], such as light emitting diodes, solar cells, transistors, diodes, holographic storage media, chemical and biological sensors, capacitors, batteries, anti-static coatings, electromagnetic shielding, anti-corrosive coatings, gas and liquid separation membranes, artificial muscles, lithography and metallization, photo-electrochromic devices, xerographic photoreceptors and all-polymer electronic circuits.

Several companies involved in commercialisation of conjugated polymer devices include, Pioneer, Dupont, Cambridge Display, Infineon, Kodak, Sanyo, Philips and TDK.

Conductive polymers have found extensive potential application in the fields of micro-sensors and micro-actuators, providing cheaper, accurate and faster alternatives to devices already on the market. Chemical sensors and biosensors with remarkable specifications have been developed and some of them are already being used in medicine for fast bio-diagnostic systems. Conductive polymer-based actuators also seem to hold great promise for the future, owing to their high strength and low energy consumption. Models of actuator response are being tested [483,654,655,495,656–658], but the relationship between input electrical energy and mechanical output has yet to be fully described.

In the sections to come, a brief description of some conductive polymers will be made, followed by a discussion of their conductivity mechanism. The focus will then shift onto applications of conductive polymers in sensor and actuator systems, including some notable examples of recent achievements.

15.1. Mechanism of polymer conductivity—role of doping

Polymers that have been studied extensively since the early 1980s include polypyrrole, polythiophene (and various polythiophene derivatives) and polyaniline. Polyacetylene, the original conductive polymer, exhibits the highest crystallinity of this group, but it was not the first conductive polymer to reach commercial production. The reason for this is its sensitivity both to atmospheric oxygen and to humidity. Polypyrrrole and polythiophene, however, differ from polyacetylene most notably in that they may be synthesized directly in the doped form and they are very stable in air.

In intrinsically conductive polymers, conductivity is attributed to delocalization of $\pi$-bonded electrons, in conjugated double bonds, over the polymeric backbone. These materials also exhibit unusual electronic properties, such as low energy optical transitions, low ionization potentials and high electron affinities [659]. Fig. 75 shows the structure of some conductive polymers.

Steric (conformational) factors and charge interactions can easily limit the degree of delocalization, leading to the formation of energy gaps analogous to conventional semiconductors such as Si or Ge. To become electrically conductive, the polymers have to be disturbed—either by removing electrons from (oxidation), or by inserting electrons into (reduction), the material. This process is known as doping. Depending on the dopant, conductive polymers exhibit

---

31 Carmen Moldovan, Rodica Iosub
either p-type or n-type conductivity. For example, treatment with halogens has been termed ‘doping’ by analogy with the doping of semiconductors. During this process, an organic polymer, either an insulator or semiconductor having a small conductivity, typically in the range $10^{-10}$ to $10^{-5}$ S cm$^{-1}$ is converted into a polymer which is in the ‘metallic’ conductive regime ($\sim 1$–$10^4$ S cm$^{-1}$). Conductivities are relatively low in comparison to metals at around $10^4$ S cm$^{-1}$, but this is enough for many practical purposes.

Dopant ions are generally introduced into the polymer system during chemical or electro-chemical polymerization and they play an important role in balancing the charge distribution within the polymer. The dopant ion influences the distribution of positive charge on the polymer backbone and, when constrained over a small area, this can cause additional closely spaced electronic levels to form within the band gap. Small lattice distortions occur locally so that conformational changes as well as charge inequalities are induced in the polymer backbone. The net effect is that the oxidation state changes and the equivalent of free radicals can be formed. A charged site interacting with a free radical forms a polaron. These are normally highly unstable but they can be stabilized by further oxidation of the polymer to form a bipolaron. Polarons are radical cations or anions that are generated during the doping process, but if the doping level is sufficiently high, a radical dication (a spinless defect) can be generated on the same polymer segment [660]. This is termed a bi-polaron. These defects lie within the band gap of the material and it is these defects that make conductive polymers interesting for chemical sensing.

The introduction of dopants and the creation of bipolarons influence the electrical conduction mechanism. It can be assumed that the attraction of an electron in one repeat unit of the polymer backbone to the nuclei in neighboring units leads to carrier delocalization along the polymer chain and to charge carrier mobility. This is further extended into three dimensions through inter-chain electron transfer. Primary dopants are used to change the electronic, optical, magnetic and/or structural properties of the polymer and this is accompanied by a large increase in conductivity. Secondary dopants also exist. These are substances that, when applied to a primary-doped polymer, induce further changes in its properties. The interactions of such molecules with the polymer, whether by hydrogen bonding, Van der Waals forces or even covalent interaction, modulate its electronic, optical, or magnetic properties. These changes can be sensed, for example, using external circuitry. Hence, the polymers can act as chemical transducers and they can be tailored to detect particular substances. The precise interactions of an external molecule with the polymer can modulate electronic coupling in different ways, but typically they lead to changes in conformation and charge transfer. The molecule being sensed can be thought of as a secondary dopant, where a charge transfer complex is formed between the polymer and the secondary dopant. If the forces involved are low, then the modulation of the electrical properties of the dopant can be reversible [661].
15.2. Conductive polymeric materials—examples

The science of conductive polymers has advanced rapidly in recent years. The most commonly used conductive polymers and their derivates will be discussed.

15.2.1. Polypyrrole

Polypyrrole has excellent electrical, thermal and mechanical properties and it can be synthesized relatively straightforward. The response of a polypyrrole film is determined by the electrochemical deposition conditions used and, in particular, the choice of the electrolyte/solvent system. Extensive studies of polypyrrole indicate that the main charge carriers are spinless bipolarons formed by the combination of two polarons [662]. In the fully oxidized, or doped state, there is on average about one positive charge for every three or four monomer units in the polymer chain. These counter-anions make up a significant portion of the polymer and they play a major role in determining the properties of the polymer [662].

In the case of pyrrole and its derivatives, coupling in the alpha or 2-position produces a polymer with extended conjugation [662]. Partial oxidation of the polymer chains induces electrical conductivity because the resulting positive charge carriers (polarons and bipolarons) are mobile along the chains.

The most commonly applied method for the deposition of polypyrrole onto sensor structures is electrochemical polymerization. Polypyrrole film can be generated from either aqueous or non-aqueous solutions, but it is thinner and more homogeneous when produced in an aqueous environment.

15.2.2. Polyaniline

English and German chemists described the polyaniline structure at the end of the 19th century but it has recently come back into prominence [663,664]. Although, Jozefowicz and co-workers noticed in 1968 a considerable raise in the electric conductivity of polyaniline, it was only after 1980 that its physical, chemical and electrical characterization became a subject of intense study. Polyaniline became the first dispersible inherently conductive polymer and it has been commercialized by Zipperling Kessler and Co.

Polyaniline was chemically synthesized by oxidation of aniline monomer with ammonium peroxydisulfate, doped with different acids and dispersed in water, alcohol or xylene. Polyaniline can be switched very quickly between insulating and conductive forms when exposed to an acidic or alkaline environment. The interest in polyaniline is also motivated by the fact that the monomer is cheap and readily accessible, chemical and electrochemical polymerization techniques are relatively simple and undemanding and the polymer is stable in ambient conditions. Chemical polymerization of aniline can also be performed in water or in various organic solvents, in acid, neutral or basic environments, using chemical oxidants.

Polyaniline has one further advantage in it can be doped into its conductive state using various protonic, anorganic and organic acids. If the organic acid has a macromolecular structure a conductive compound can be obtained. Polyaniline can be made conductive by a protein acid-treatment that does not involve irreversible protonation of the polymer backbone [665]. Protonic doping of polyaniline has also been achieved using inorganic acids such as HCl, H₂SO₄, HNO₃ and organic acids such as CF₃SO₃H and CF₃COOH [666].

15.2.3. Polythiophene

This compound has recently gained special attention as a conductive polymer because of its thermal and oxidative stability, its optical nonlinear properties and its high electrical conductivity in doped state. Because the oxidation potential of thiophene is greater than that of pyrrole, the number of chemical oxidants that can be used for polymerization is much smaller.

As in the case of other polyheterocycles, synthesis of polythiophene may be performed both chemically and electrochemically. Although, chemical synthesis is a method that yields a larger quantity of polymer, electrochemical polymerization of heterocycles is more effective, because in this case the final polymer is free of impurities such as catalyzers and metallic salts. It can be deposited electrochemically as adherent films whose thickness can be controlled and the electro-polymerization may be performed in either potentiostatic or galvanostatic conditions. The substituent(s) and the side group(s) attached to thiophene allow more extensive electron delocalization, which helps to lower the band gap energy and reduces the energy required for the polymer to be charged. The ideal structure is made
up by coupling of the thiophenic rings in α,α' or the 2,5 positions. Substituents may alter the properties of polythiophene in two ways:

(a) By introducing photo-conductive, chiral or liquid crystal, groups which may subsequently complex with various ions or organic compounds.
(b) The volume of the substituent may also determine the torsion angle between the thienyl groups, hence the length of the conjugation and the position of the absorption π-π.

Polythiophenes having alkoxy groups as substituents are very interesting because they promote solubility and increased hydrophilic character. They can complex with cations to make chemical sensors. Moreover, when the alkoxy group is chiral, the polymer can form complexes with enantiomers out of mixtures of optical isomers. Polythiophenes with macrocyclic ether groups (crown ethers) are very important because of the possibility of cation complexion and self-assembly. These macrocycles may be attached:

(a) By a spacer in position 3 of thiophene.
(b) By monomers containing two thiophenic groups, connected in positions 3,3' by poly-etheric groups or thiophene with polyetheric groups introduced in position 3.

Polythiophene derivatives can be made highly conductive by treatment with FeCl₃, AuCl₃, I₂ or H₂SO₄. Oxidized films exposed to ambient light and air showed a loss in conductivity. The lifetime of the conductivity was dependent upon the stability of the oxidizing agent used.

15.2.4. Polysiloxane

Polysiloxane [Si(R)₂-O]ₙ is the most studied and most important of the inorganic polymers. Due to its very strong bonding between silicon and oxygen, polysiloxane is generally extremely stable. The characteristics of the polymer depend on the reaction conditions during synthesis. Basic catalysts and high temperatures lead to the preferential formation of linear polymers with high molecular mass. Polysiloxane polymer hosts with branched and crown ether side chains exhibit high ionic conductivities [667]. Most recently work has split into two directions: respectively, lower viscosity liquid siloxanes and siloxanes that are incorporated into freestanding gels. Polymethylsiloxane polymers with an oligo(ethylene glycol)methyl ether side chain are lower molecular weight hosts and liquids. The doped electrolytes have ionic conductivities between 10⁻⁵ and 10⁻⁴ S cm⁻¹.

15.2.5. Polyphthalocyanine

Phthalocyanine is a macrocyclic compound having an alternating nitrogen atom-carbon atom ring structure, Fig. 76 [668]. It was probably discovered by accident in 1907, as a by-product during synthesis of o-cyanobenzamide, but it was not until almost 20 years later that a patent was filed describing a manufacturing process.

The molecule is able to coordinate hydrogen and metal cations in its center by bonding to the four isoindole nitrogen atoms. The central atoms are able to carry additional ligands. Most elements are able to coordinate to the phthalocyanine macrocycle and therefore, a variety of useful phthalocyanine complexes can exist. The structure of a phthalocyanine molecule is closely related to that of the naturally occurring porphyrin systems. Complex investigation
of electric and gas-responsive properties of a series of synthesized metal-free and metal-containing oligo and polyphthalocyanines (OPc and PPc, respectively) has been reported.

15.2.6. Fullerene

The 1996 Nobel Prize in Chemistry was awarded for the discovery of fullerenes, which are one of only four types of naturally occurring forms of carbon. The fullerenes are molecules composed entirely of carbon, taking the form of a hollow sphere, an ellipsoid or a tube. Spherical fullerenes are sometimes called buckyballs, while cylindrical fullerenes are called buckyballs or nanotubes. Fullerenes are similar in structure to graphite, which is composed of a sheet of linked hexagonal rings, but they contain pentagonal (or sometimes heptagonal) rings that prevent the sheet from being planar. The smallest fullerene is C_{60} (buckminsterfullerene) and it is also the most common.

The buckminsterfullerene molecule, a closed cage structure molecule with carbon network, is shown in Fig. 77 [669]. The structure of C_{60} is that of a truncated icosahedron, which resembles a round soccer ball of the type made of hexagons and pentagons, with a carbon atom at the corners of each hexagon and a bond along each edge.

Novel electrical and optical characteristics have been observed in conductive polymers doped with molecular dopants such as fullerenes (C_{60}, C_{70}). The C_{60}/conductive polymer junction can be interpreted as donor–acceptor (D–A) type photocell rather than the conventional p–n junction common in organic semiconductors [670].

15.3. Applications of conductive polymers in sensors and actuators

The development of biological and chemical sensors based on conductive polymers has been investigated extensively during the past ten years. Essentially, they can function in two different ways: as conductive materials for the construction of various devices and as selective layers in chemical sensing applications [671]. In both cases, the interaction with ambient gases is an important factor to consider. On one hand, it may compromise the performance of a device based on conductive polymers. Conversely, it may be beneficial in a sensor. Conductivity has been the primary focus of recent interest. Work function, which is related to conductivity but in principle a different property, has received only scant attention.

The advantages in using conductive polymers as sensing materials over conventionally used materials [661] include: (a) the availability of a diverse range of monomer types and synthetic monomer analogues; (b) electrochemical preparation readily allowing mass production and miniaturization of the sensors; (c) the ease with which biomaterials, such as enzymes, antibodies and whole cells, may be incorporated in the polymer; (d) the ability to change the oxidation state of the polymer after deposition and thereby tailor the sensing characteristics of the film; (e) the ability to obtain a reversible response at ambient temperature.

15.3.1. Sensors

A variety of conductive polymers have been evaluated using microelectronic devices, such as chemoresistors (interdigitated array transducers), quartz crystal microbalances (QCMs) and field-effect transistors (FETs). Kunugi et al. [672] utilized a specially modified QCM for making electrical and microgravimetric measurements of the uptake of alcohols onto polypyrrole thin films. Josowicz and Janata investigated the measurement of work-function changes using a polypyrrole-coated FET for the detection of lower aliphatic alcohols [673]. Several companies, including
Neotronics [674] and AromaScan [675], manufacture “electronic noses” comprising arrays of chemoresistor-based conductive polymer sensors.

Conductive polymers are an effective medium for chemical sensing, based upon the electronic changes that are induced in the polymeric films due to gas-phase adsorption. The conductivity changes are commonly attributed to the interaction of analytes with either the polymer backbone itself or with dopant molecules incorporated within the film, thereby modulating the mobility and/or the number of free charge carriers available [676].

Conductive polymer sensors can be applied in a number of different modes [677]:

- pH-based (change in pH);
- conductometric mode (change is conductivity);
- amperometric mode (monitoring the variation in current);
- potentiometric mode (change of open circuit potential).

Gas sensors are often operated in conductometric mode, whereby changes in the conductivity or resistivity are measured while exposing the polymer to the gas molecules of interest (electronic noses).

To-date, most research on biosensors has focused on devices that consist of a metal layer to conduct electrons as the signal, an enzyme or antibody to sense the presence of a particular analyte, and a membrane to immobilize the enzyme and also possibly aid in transducing the signal to the electrode. Conductive polymers are used in biosensors in three main areas: signal detection, transmission of signal to a measuring electrode or response element and controlling the feedback response to the signal. They are characterised by their exceptional versatility and at the present moment in time they promise to revolutionize chemical sensing.

15.3.2. Chemical microsensors

Two types of sensor, developed at Sandia National Laboratory, were tested at the Nevada Test Site as part of the Advanced Monitoring Systems Initiative program in late 2002. These were a chemo-resistive sensor and a surface-acoustic-wave (SAW) sensor. Both sensors utilize the adsorption of chemicals onto polymer films to produce a change in an electrical signal, which is recorded and calibrated. The results showed that both sensors could be operated remotely and continuously for long-term monitoring applications using commercial data-acquisition systems and custom-designed packaging. Both the chemo-resistive and the SAW sensors experienced drift in the signal and the readings were impacted by fluctuations in temperature and humidity. The chemoresistor, however, showed that exposure to large concentrations of contaminants overwhelmed the fluctuations caused by temperature and humidity variations. It was also demonstrated that the chemoresistive sensor exhibited better stability and sensitivity than the SAW sensor for the conditions and analytes that were tested and this was contrary to initial theoretical predictions [678]. The chemoresistor sensors used in the field tests at the Nevada Test Site detect volatile organic compound (VOC) vapours with conductive polymer films deposited onto micro-fabricated circuits. The polymers used were: a 75% phenyl:25% methylpolysiloxane (OV25); an alkylaminopyridyl-substituted polysiloxane (SXPYR); a hexafluoro-2-propanol-substituted polysiloxane (SXFA); polybis(cyanopropyl)-siloxane (SXCN); poly(vinyltetradecanal) (PVTD) and poly(trifluoropropyl)methylsiloxane (OV-202). The variety of compounds that could be detected in this way included aromatic hydrocarbons (benzene), chlorinated solvents (trichloroethylene, carbon tetrachloride), aliphatic hydrocarbons (hexane, iso-octane), alcohols and ketones (acetone).

In order to fabricate this type of sensor, the chemically sensitive polymer is dissolved in a solvent and mixed with conductive carbon particles. The resulting ink is then deposited onto a solid substrate, between some platinum tracks and dried. When chemical vapour comes into contact with the surface it is absorbed by the polymer, causing it to swell. The swelling changes the electrical resistance which can be measured and recorded. The amount of swelling corresponds to the concentration of the chemical vapor in contact with the chemoresistor and the devices can be calibrated by exposing them to known concentrations of target analytes. In the Nevada tests, the SAW sensor probe was placed alongside the chemoresistor probe (Fig. 78), according to ref. [678].

The conductivity of polyaniline is highly sensitive to chemical vapours and, besides polysiloxane, this is another promising material for manufacture of chemical sensors. An example is given in ref. [679], where oriented polyaniline nanowires, with diameters on the order of 100 nm were deposited on gold electrodes. The devices showed a rapid and reversible resistance change upon exposure to NH₃ gas at concentrations as low as 0.5 ppm. The well-defined single-wire geometry allows for accurate characterization of the device response. The response times of nanowire sensors
The nanowires were deposited, using a particular scanned micro-fabricated electrospinning source which presents a general method for interfacing polymeric nanowire devices with microfabricated structures. The method could ultimately be used to make a device that has an array of wires sensitive to different chemicals. Such a device would be able to rapidly detect and analyze the composition of gases in the air [680].

Other chemical sensors based polyaniline are ammonia sensors fabricated on a dielectric substrate using microelectronic planar technology. A sensitive layer based on polyaniline is deposited onto interdigitated metallic electrodes. The principle of operation lies in the variation of the electrophysical characteristics of the film (conductivity, charge carriers concentration) that is induced by the reversible ammonia adsorption process and its chemical interactions [681]. In practice, the sensors are found to be high selective to ammonia in mixtures with different active toxic gases (nitrogen dioxide, hydrogen, methane, hydrogen sulfide, carbon monoxide, carbon dioxide). The sensitivity threshold is 0.1/0.5 ppm and the range of measurement is between 0.1 and 1000 ppm. Response time is 10–15 s and the sensor can operate at room temperature. These sensors are suitable for measurement of ammonia content in an industrial environmental monitoring system, for some agricultural applications, in medical and biological research.

15.3.3. Electronic noses

Polymer membrane sensor arrays, also called ‘electronic noses’, have been in research for several years and they could provide a simple alternative to many complex instrumental techniques, such as chromatography and ultra-violet spectroscopy, which are both expensive and time-consuming. Some successful trials of the technology have taken place at hospitals in Manchester and London to diagnose bacterial vaginosis, BV).

A company called Osmetech has designed the electronic nose customized to smell out gases emitted by the bacteria that cause BV. This ‘nose’ is as accurate as other methods and tests can be performed in very much less time. The analyzer instrument consists of an array of 48 of these conductive polymer sensor elements, made using wafer-scale technology. Arrays can be selectively tailored to match the particular application: for example, a narrow range of sensor types would be chosen to obtain very high sensitivity where only a few chemical species need to be assayed. A module containing a broad range of different sensors would be used where many different responses are needed to detect a wider range of analytes, (Fig. 79) [682].
Conductive polymers are made by ‘electropolymerisation’ of complex organic dyes—specifically, derivatives of the substances polypyrrole, polyaniline and polythiophenes.

Osmetech have discovered how to give each of these polymers different conductive behavior, through variations in chemical structure, so that it has managed to build up a proprietary library of 80–100 sensor materials, each applicable to a particular molecular type and each sensitive to variations in concentration. The responses of individual sensing elements are picked up by signal processing circuits and compiled into a response profile (or ‘scent’) that characterises the atmosphere around the sensor.

An electronic nose also has many industrial applications and important amongst these is detection of aroma. Food producers and manufacturers of toiletries use panels of people to assess the smell of their products. Naturally, human factors ensure that the panels can never be entirely objective or in fact reproducible. Electronic noses can provide objective data, and polymeric sensors are now being used to sniff out the amount of ethanol in mouthwash, monitor odour levels in sodium bicarbonate for toothpaste, check the stability of fragrances in soap tablets and to assess the smell of fixatives used in hair sprays.

Another application is in cereal quality assurance. Specialty trained human sniffers have been used for hundreds of years to detect the presence of mould and contaminants in bulk grain stores. This is particularly important in the Far East, with its large rice consumption; rice stored for a long time can acquire unpleasant tastes and smells, which significantly affect its price. The polymer sensor array can perform this sensing role and supply some consistency in assessing the quality and value of the rice.

Polymer sensor arrays are also being used to check the usability of recycled plastics, which are nowadays-popular materials for manufacturing car interiors. Such plastics can retain traces of smells from their previous uses, which are not necessarily welcome inside a car on a hot day. Sensors are available that can assess these trace odours much more rapidly than is possible using gas chromatography—and in a real vehicle environment, too, not just in the lab. [683].

15.3.4. FET type devices

Electronic devices can be made, for example, by forming tracks of PEDOT (poly-3,4-ethylenedioxythiophen) on a substrate using a solvent evaporation, line patterning, technique. Two electrodes were prepared in this way, comprising a source/drain electrode and a gate electrode [684,685] (Fig. 80).

Line Patterning is a novel process that provides a simple and cheap way to prepare patterns of conductive polymers. It exploits differences in selected physical and/or chemical properties between a substrate and insulating lines that have been printed onto it by a conventional copying or printing process. The substrate and printed lines react differently or at different rates with a solvent (or vapor) to which they have been exposed and, hence, the conductive polymer can be selectively deposited as a film after evaporation of the solvent. Line patterning has the advantages that no photolithography is involved; no printing of conductive polymer is involved; it uses only, for example, a standard office laser printer, which is not modified in any way. Commercially available flexible, transparent plastic or paper substrates can be used and solutions of conductive or non-conductive polymers are commercially available.
15.3.5. Biosensors

An enzyme electrode is a combination of any electrochemical probe (amperometric, potentiometric or conductometric) with a thin layer (10–200 nm) of immobilized enzyme. In these devices, the function of the enzyme is providing selectivity by virtue of its biological affinity for a particular substrate molecule. For example, an enzyme is capable of catalyzing a particular reaction of a given substrate even though other isomers of that substrate or similar substrates may be present. The progress of the enzyme reaction (which is in turn related to the concentration of the analyte) is monitored by either the rate of formation of product or the disappearance of a reactant.

The method of analysis used depends on several properties of the enzyme such as: the existence of redox active groups on the enzyme, the electrical properties of the products of the biochemical reaction, the electrical characteristics of the substrates or cofactors, the required speed of response and ultimately the intended application of the sensor. The first three of these criteria will depend largely on the system under investigation. The latter three depend on the requirements and application of the sensor under consideration. If the enzyme does not contain any redox groups, then the method of analysis will be restricted to monitoring either the release of products or the consumption of substrate by their reaction at the transducing electrode. The current produced can then be related directly to the concentration of analyte [686].

The glucose sensor is one of the most successful commercial biosensors, used to monitoring blood sugar levels for diabetics. An enzyme deposited on an electrode surface acts as the bio-recognition element to identify glucose molecules. When an enzyme recognizes the glucose molecules, it acts as a catalyst to produce gluconic acid and hydrogen peroxide using oxygen from the air. The electron transfer due to hydrogen peroxide/oxygen coupling is easily detectable and the electron flow is proportional to the number of glucose molecules that are present in blood.

A new type of glucose sensor uses the hydrogen-specific gas sensing capability of single walled carbon nanotubes (SWCNT) assembled on a micro-electrode. The conductivity of SWCNT’s effectively changes due to adsorption of hydrogen owing to a shift in the energy band structure. Such devices could be minimally invasive and possibly non-invasive, relying only upon a small sample quantity of blood or other body fluid. It could also be highly sensitive, very fast immune to false positives and it would not require any calibration. It has been demonstrated that a signal-readout is proportional to glucose levels within a clinically significant range (Fig. 81) [687]. Each device requires a few thousand-carbon nanotubes on an area of a ~1 mm × 1 mm in size). Only electrical resistance needs to be measured, and the sensor can be easily operated without specific training.

Microtubule sensors for glucose, urea and triglyceride have been fabricated based on poly (styrene sulfonate)-polyaniline (PSS-PANI) composites. These were synthesized within the pores of track-etched polycarbonate membranes and arrays were fabricated by immobilization of three different sets of enzymes on three closely spaced devices. The sensors for urea and tri-glyceride were found to have a higher linear range of response, better sensitivity, improved multiple-use capability and faster response time in comparison to potentiometric and amperometric sensors based on polyaniline. Response was found to be free from cross-interference when a sample containing a mixture of the above analytes was analyzed in a single measurement [688].

15.3.6. Actuators

Conductive polymer actuators are rapidly advancing from promising laboratory research [487,689–694] towards full-fledged engineering materials [695,654]. Most studies to-date have investigated the contractile properties of either

![Fig. 81. The schematic of the glucose sensor [687].](image-url)
of two conductive polymers, polypyrrole and polyaniline. These conductive polymer actuators are of technological
interest due to their low operating voltages, high forces, moderate strain and controllability. The operating voltage
range is typically less than 1 V, with short bursts of up to 10 V being used to increase power [495]. Such voltages are
readily provided by batteries and of course this is of particular value for micro-scale devices applications [483]. The
fact that displacement is proportional to charge transfer in conductive polymer actuators makes them relatively easy to
control, giving them a potential advantage over shape memory alloys [656]. Unlike mammalian muscle and
electromagnetic motors, conductive polymers do not expend significant energy when holding a load in place, greatly
reducing losses in static applications such as switches and valves [495]. Conductive polymers show great potential as
electro-mechanical actuator materials although efficiency currently tends to be low at around 1%.

Conductive polymers actuate via the reversible counter-ion insertion and expulsion that occurs during redox
cycling. Significant volume changes occur through oxidation and reduction reactions at corresponding electrodes
through exchanges of ions with an electrolyte. Electrodes are commonly fabricated from polypyrrole (Ppy) or
polyaniline (PAN), or PAN doped with HCl. Polypyrrole actuators in particular can generate a strain of 1–3% under
electro-chemical excitation and the forces generated are 100–1000 times greater than skeletal muscle. They require
low voltage for actuation (1 V or less), are biocompatible and operate ideally in liquid electrolytes (including bio-
fluids). Such features make them very promising for biomedical applications.

Responsive drug delivery applications of polymer actuators have been studied at UC Irvine. In this case drug
reservoirs were incorporated with small polymer valves, biosensors control the opening and closing of these valves and
the required amount of drug is released [696]. An example of an actuator made using conductive polymers is shown in
Fig. 82.

The applications for conductive polymer actuators include microtweezers, microvalves, micropositioners for
microscopic optical elements and actuators for micro-mechanical sorting (such as the sorting of biological cells).
Furthermore, these materials offer the possibility of creating devices at the nano-scale. Nano-electromechanical
systems (NEMS) devices and machines are an extension of present-day micro-machines and micro-actuators into the
nano-scale domain. Notable achievements in this field of activity are protein motors, capable of linear or rotary
motion, DNA and active devices such as nanowires, switches, motors and tweezers [697,698].

15.4. Conclusions

Although implementation of these materials is still at its beginning, it is becoming clear that they will permit
important (if not radical) changes in future electronic systems. Conjugated polymers are expected to play, for example,
a crucial role in the fabrication of molecular electronic circuits. In the future, the dimensions of electronic circuits
could be reduced well below the current minimum pattern dimensions used in advanced integrated circuits on silicon
wafers, perhaps reaching as small as 1 nm. A reduction in circuit size of this magnitude could increase the speed and
the dynamic memory capacity of computers by several orders of magnitude (~10^6). Sensitive detectors have been
made that can detect low ppm levels of toxic gases. These can be inexpensive, compact, simple to operate and they can
potentially be incorporated within a fabric-based sensor system.

A number of enzymes have been covalently attached or integrated into conductive polymers as dopants.
Interestingly, enzyme activity is dependent on the potential applied to the conductive polymer, and this can be used to
modulate its behaviour. In a similar way, antibodies can be incorporated into conductive polymers as dopants.
Application of a potential to the polymer can change the conformational characteristics of the antibody so that reversible affinity biosensors can be constructed.

Conductive polymer actuators require voltages in the range of 1–5 V. Relatively high mechanical energy densities of over 20 J/cm$^3$ are attained with these materials, however, they possess low efficiencies at levels of 1%.

16. Polyimides

Polyimides (PI) are part of a group of high-temperature polymers that have been applied in microelectronics and microelectromechanical systems (MEMS). They constitute an important class of materials because of their low dielectric constants, high breakdown voltage, good planarization, wear resistance, radiation resistance, inertness to solvents, good adhesion properties, low thermal expansion, long-term stability and excellent mechanical properties.

Due to their outstanding features, polyimides, photosensitive and non-photosensitive, have found many applications both in integrated circuitry (IC) and microfabrication. In these areas, PI films are being used as structural materials, passivation layer, interlayer dielectric or protection layer. In addition, an important commercial application of PI is the construction of multilayer flexible circuits used in the realisation of multi chip modules (MCMs). On the other hand, photosensitive PI precursors, called polyamic acids, can also be used as new photoresists with promising sensor and microfabrication applications.

PI films are supplied either as films (obtained by lamination) or as solutions that can be applied by spin coating, afterwards curing over 275 °C to convert to PI.

16.1. Properties of polyimides

PI film has a unique combination of properties that make it ideal for a variety of applications in many different industries. PI film can maintain its excellent balance of electrical, thermal, mechanical, physical and chemical properties over a wide range of temperatures. In MEMS structures and flexible circuits, the most widely used are 25–50 μm thick flexible Kapton from DuPont [699] and Upilex from Ube Industries [700] PI films. Han et al. in ref [701] gathers comparative tables that show the mechanical, chemical and electrical properties of Kapton and Upilex PI films. These tables can be extremely useful in order to make a good choice of materials for applications. Both the Kapton and Upilex PI films possess very strong mechanical properties, sufficient chemical resistance, and excellent thermal and electrical characteristics. However, Upilex PI film is relatively stronger mechanically, and shows higher chemical resistance and dimensional stability over a wide range of temperatures than Kapton PI film, while Kapton is stronger in folding endurance, and shows higher electrical resistance and lower thermal conductivity than Upilex PI film.

16.2. Processing of polyimides

The conventional methods to pattern PI films are by laser, NC drilling or punching, wet-etching (achieved using chemicals) or dry-etching (using plasma or reactive-ion etching) and photolithography, in the case of using photosensitive solutions. In principle, the laser, mechanical and wet etching processes produce coarse features whereas the dry etching and the photosensitive polyimides allow for finer features. Fig. 83 shows the established methods for wet etch, dry etch and the use of photosensitive polyimide.

Both wet-etch and photo-definable polyimides are patterned prior to final cure. Whereas, dry etch processing requires the polyimide layer to be fully cured prior to the etching process.

16.2.1. Wet etch patterning

Wet etch patterning uses wet chemistry in conjunction with a liquid photoresist to define the desired pattern in the polyimide layer. Wet etch processing is typically used to pattern coarse features such as bond pads or large vias. An aspect ratio of 1–5 can be achieved reliably.

The base process involves spin coating and partial curing of the polyimide layer, using one or more in-line hot plates or sometimes a convection oven. A layer of positive photoresist is deposited, baked and imaged. The photoresist is then

32 Estefania Abad
developed with a standard developer. A TMAH-based developer can simultaneously wet etch the underlying polyimide in the imaged areas. Otherwise alkaline solutions, like the TPE 300 from Toray [701] or the Kapton polyimide film etchant from Transene [702] can be used. After develop/etch and a water rinse, the photoresist is stripped using a liquid photoresist stripper. The patterned polyimide wafer is then fully cured to complete the imidization process and remove residual solvent.

16.2.2. Dry etch patterning

Dry etch patterning is typically used when an aspect ratio greater than 1 to 1 is required for very fine feature sizes. The process begins with the application and full cure of the polyimide layer. After curing, a dry etch mask is applied over the polymer layer. Thin films of aluminium or CVD oxide are typically used for this purpose. Photoresist is applied and imaged over the etch mask to define the desired pattern. The wafer or substrate is then placed in a plasma or reactive-ion etcher, typically of parallel plate configuration. Process parameters and gas mixture are then optimised to pattern the etch mask. The gas mixture is switched to an oxygen/CF$_4$ blend and parameters are set to etch the underlying polyimide and remove the photoresist. The etch mask is then stripped by switching back the etch parameters or by using a wet process.

16.2.3. Photodefinable polyimides

The use of photosensitive polyimides simplifies the process complexity by reducing the number of steps as shown in Fig. 83, as well as improving the sidewall profile. Photosensitive polyimides permit the patterning of relatively fine features. An aspect ratio of 1 to 1 can be achieved in fully cured films. HD Microsystems is a global supplier of liquid polyimide materials [703].

The basic process involves the spin coating of the polyimide and a drying step, using one or more in-line hot plates or a convection oven. The polyimide layer is then exposed on a standard I or G line lithography tool. A negative tone
photo mask is usually required since most photodefinable polyimides are negative acting. However, positive aqueous soluble photosensitive polyimides have been formulated recently [704]. After imaging, the wafer is developed on a traditional track line and rinsed. Finally, the polyimide layer is cured, both to imidize the film and to remove the photo package.

16.2.4. Laser ablation

Laser ablation also begins with a fully cured film. The wall profile of the vias formed with a laser is V-shaped and it is not aspect-ratio limited. The specific patterning process depends upon the laser tool configuration. Most of the reported works on laser processing of polyimides, in particular for flex interconnect applications, use excimer or Nd:YAG laser [705]. However, kapton processing using other types such as CO₂ or femto-second laser is also possible. The advantage of using femto-second laser equipment is that it combines the use of UV light along with the femto-second pulses, to avoid the creation of heat affected zones on the polymer [706].

The process starts by transferring a DXF file to the code of the equipment. The process parameters that need to be optimised are power, focal length, advance (which determines the line width) and the thickness of material to be ablated. When these parameters have been established, the system can operate unattended. Fig. 84 shows blind via holes 760 µm diameter and through holes fabricated in the same femtosecond laser ablation process. The material employed is Pyralux AP 8525R double-sided, copper-clad laminate from DuPont. It is an adhesiveless laminate for flexible printed circuit applications. The kapton has a thickness of 50 µm and the copper layer has a thickness of 18 µm on each side.

16.3. Polyimide applications

16.3.1. High density interconnection flexible substrates

Sophisticated high density interconnection (HDI) flexible substrates are experiencing a strong demand for advanced applications such as implantable medical devices [707,708], as well as for new products in the printing and telecommunications industry.

Flexible substrates offer some unique advantages when compared to their rigid counterparts and in particular they are lighter, thinner and in some applications they are adaptable to more flexible packages. To create such a HDI flex circuit a multilayer construction is needed. Either thin film or laminate technology can be used. Thin film technology is based on polyimide spin coating, patterning, metal deposition and patterning of successive layers. Laminate technology is used to make polyimide sandwich structures with layers of copper. Each layer has to be individually fabricated, assembled by multilayer construction (alignment and bonding) and Z interconnected through vias. The main advantage of using the laminated approach is its lower cost in comparison to the thin film approach, which is a relatively immature technology.

16.3.2. MEMS devices

Polyimide is a useful material for MEMS fabrication. The use of polyimides in this field is quite recent but it includes some interesting applications as it can be seen from the following examples.

Fig. 84. Via holes blind via holes of 760 µm diameter and windows fabricated by femtosecond laser ablation in a Pyralux composite from DuPont.
16.3.2.1. Humidity sensors. Polyimide films have been recently introduced for the fabrication of capacitive humidity micro sensors [709–711]. The principle of a humidity capacitive sensor lies on the increase of the dielectric permittivity of a sensitive layer due to the moisture absorption or in the change of capacitance due to moisture dependent bending of a movable electrode. Polyimide films can be used either as a sensing layer or as a vapour absorbent polymer coating for the movable electrode. The sensitivity of this polyimide layer can be amplified by patterning the layer to increase the active surface area or by treating the surface in an oxygen plasma to raise the roughness. Polyimide is a polymer particularly suitable for this application thanks to its high permeability to water, its great chemical stability and its full compatibility with IC process.

16.3.2.2. Implantable flexible electrodes. Neural prostheses are systems that partially substitute for the neural functions of the body that have been lost after traumatic lesions or neurological disorders. This loss can be partially replaced by implantable flexible electrodes for functional electrical stimulation of the remaining nerves or muscles. The most well-known examples of this type of devices are cardiac pacemakers and cochlear implants. These types of systems are usually a combination of flexible micromachined substrates with monolithically integrated electrodes, hybrid assemblies of electronics and encapsulation materials. Polyimide substrates are frequently used because they are flexible, light and can be easily patterned [712].

16.3.2.3. Flexible tactile sensor skins. Tactile sensing is an area of MEMS research that has potential applications in robotics, medicine and industrial automation. A tactile skin is a two-dimensional array of flexible sensor substrate with a high density of sensors of the following types: a hardness sensor, a temperature sensor, a sensor for thermal conductivity and sensor for surface roughness measurement. The realization of multimodal tactile skin involves the application of MEMS technology on polymer materials such a polyimide. Fig. 85 shows a novel fabrication process of a skin containing a two dimensional array of tactile sensors using polyimide and metal strain gauges developed by Engel et al. [713]. The result is a flexible, robust, monolithical polyimide-based sensor with embedded metal-film strain gauges and interconnects to detect force distribution. Sensor fabrication, on oxidized silicon wafers or Pyrex wafers with recessed regions (Fig. 85(a)–(c)), consists of the deposition, patterning and curing of three layers of photosensitive polyimide with embedded metal thin films patterned by lift off (Fig. 85(d)–(h)). Finally, the polyimide devices are released from the wafers Fig. 85(i) and packaged on flexible copper-clad substrates.

16.3.2.4. Piezoelectric polyimide MEMS process. Recently, a series of novel high-temperature piezoelectric polyimides have been synthesized and evaluated for potential use in MEMS devices [714]. Typical piezoelectric materials are crystalline ceramics that are brittle. Alternatively, polyvinylidenefluoride (PVDF) has attained a widespread use as piezoelectric material in low cost flexible structures. However, the temperature range over which the piezoelectric properties of PVDF are maintained is limited to less than 80 °C. For applications at higher temperatures, as, for example, in the aerospace field, the introduction of this type of piezoelectric polyimides is of great interest. In order to demonstrate sensor and actuator utility of piezoelectric polyimides a low temperature (<120 °C) surface

![Fig. 85. Schematic diagram of the fabrication process of a single taxel [713].](image-url)
miocromachined process has been developed by Atkinson et al. [715]. This process consists of conventional lithographic and metallisation steps and uses a sacrificial layer of photoresist.

16.3.2.5. Polyimide joints. Creation of 3-D structures is a very important issue in the field of MEMS and several different techniques have been proposed. However, many of them have drawbacks. Eberfors et al. [716] developed a novel and simple technology for making robust, small radius micro-joints (or micro-hinges) based on thermal shrinkage of polyimide in V-grooves. The structures can be used in both static-mode, for 3-D sensor applications and dynamic mode for use as actuators. Polyimide is an excellent micro-joint material because of its flexibility, high thermal shrinkage and high thermal expansion.

16.3.2.6. Polyimide sacrificial layers for surface micromachining. Compatibility of surface micromachining with conventional IC processing is essential for MEMS devices. Several post-process approaches have been developed in recent years, including some that use polyimide as a sacrificial layer [717]. Micromachining post processes require deposition and patterning of mechanical and sacrificial layers at temperatures below 400 °C. To avoid stiction problems during the release of the sacrificial layer, dry steps as oxygen plasma etching have been used to eliminate the polyimide.

In addition to these illustrations of polyimide applications, many other examples exist, such as a polyimide membrane with ZnO piezoelectric thin film pressure transducer, for use as a differential pressure liquid flow sensor [718]; high-resolution polyimide-based double-layered coils on flexible Kapton support for miniaturized power inductors [719]; or the mould transfer fabrication of polyimide SPM probes [720]. All these examples reveal the exceptionally wide field of applications that polyimides can cover.

17. Structural polymers

In recent years, silicon has been the dominant mechanical material for micro-electromechanical systems (MEMS) fabrication [721]. However, other materials such as glass, ceramics and metals are equally useful at this scale and they have also been used as structural or functional materials with great success [722,723]. The term ‘structural’ usually denotes those materials that are used solely for mechanical purposes. Moving parts in mechanisms also fall into this category. Despite the great benefits of using traditional materials such as silicon, it is not always true that they are the best choice for structural applications in micro-systems. In particular, their high modulus and related stiffness restricts the development of micro-sensors and actuators that would gain in sensitivity and performance, or could be further miniaturized by having moving parts with reduced stiffness [724–726]. Soft materials, such as synthetic polymers, are also excellent candidates for structural use in micro-scale sensors and actuators, since they are available with tailored moduli in the range <1 MPa up to several GPa [727,728]. For example, the modulus of chemically cross-linked elastomers may be tuned over one or even several orders of magnitude by controlling the amount of cross-linking between the polymer chains [729–732]. In addition to the wide range of mechanical properties available, the transparency of many polymers is an advantage if visual control is necessary, for example in micro assembly processes [724]. Furthermore, many polymers offer good biocompatibility [733], which is either necessary or desirable for some biomedical applications such as actuated micro valves in micro-fluidic systems [734,735].

On a macroscopic scale, synthetic structural polymers have become the most widely used material in terms of volume production [736–738]. Here, the main requirements for a structural polymer are mechanical stability and low material costs [739]. Nowadays, the so-called ‘commodity’ or ‘standard’ polymers such as high and low-density polyethylene (HDPE, LDPE/LDPE), polypropylene (PP), polystyrene (PS) or polyvinylchloride (PVC) have gained over 80% share of the industrial structural polymer market [738]. On a microscopic scale, the requirements for structural polymers are somewhat different. Apart from good mechanical properties, their ease of processing plays an important role in the materials selection for micro-scale devices. Recently, significant advances have been made in lithographic technologies for polymers, which make use of their low softening temperature or their sensitivity to radiation. In particular, photosensitive polymers such as SU-8 [740] have attracted great attention for the fabrication of
microstructures with high aspect ratios [741–744]. Micro-molding [745–747], rapid prototyping [748] and the so-called “soft-lithographic” techniques [749–756] such as hot-embossing have been developed and refined so that they are now well suited for micro and nano-fabrication of non-photosensitive, thermoplastic polymeric components. The term ‘soft lithography’ has been adopted from Xia and Whitesides [749] for non-photolithographic micro-machining techniques that use soft polymer materials as the stamp, mould or mask to generate micro-patterns and microstructures. In addition to soft lithographic techniques, most polymers can be structured by established surface micromachining techniques such as dry etching [483,757].

Release of free-standing micro-structures can be achieved by sacrificial-layer techniques, whereby a sacrificial layer is deposited below the structural element that will be released. Subsequently, this sacrificial layer is dissolved or removed by selective etching. The correct selection of the sacrificial layer material needs great care, since ideally the release step should not affect the structural polymer. This can be a challenge [483,758], particularly for polymers with poor chemical resistance. In this case, water-soluble sacrificial polymer layers might present a potential alternative [759].

For moving structures there is a challenge related to the fact that polymers, particularly thermoplastics, can exhibit enhanced viscoelastic behavior [760–762]. This has consequences for resonant applications and for the long-term stability of dynamic microstructures. The viscoelastic nature of polymers can cause high energy dissipation by enhanced damping, or by creep deformation under load. Therefore, polymers intended for structural use in micro-scale sensors and actuators require a thorough evaluation and characterization of their viscoelastic behavior.

In the following paragraphs, those polymers will be discussed that have been already used as structural materials in micro-scale sensors and actuators, which are currently under investigation or which have high potential for applications in MEMS. The reader is referred to references [763–766] for further background information.

17.1. Selection of structural polymers for micro-scale devices

A summary of important material properties of polymers in this section can be found in Table 11.

17.1.1. Thermosets

Thermosetting plastics (thermosets) are polymeric materials that form cross-linked networks on curing. Curing is an energy-activated process that can in different cases be initiated by heating, through a chemical reaction, or by irradiation. Low molecular-weight precursor materials can be used that can be deposited as very smooth layers of varying thickness by solution spin casting [767] and subsequent curing. Generally, this yields materials that have a high modulus and high thermal stability in comparison to most non-thermosetting polymers. The viscoelastic nature of polymers can cause high energy dissipation by enhanced damping, or by creep deformation under load. Therefore, polymers intended for structural use in micro-scale sensors and actuators require a thorough evaluation and characterization of their viscoelastic behavior.

SU-8, is a low-cost, negative photo-resist developed by IBM in 1982 [769]. It is based on a multi-functional glycidyl ether derivate of bisphenol-A novolac. EPON® SU-8 from Shell Chemical currently provides the highest functionality that is commercially available [770]. This low-molecular weight resin can be dissolved in a variety of organic solvents up to very high concentrations (85 wt%) and it can be photo-sensitized with triaryl sulfonium salt [740]. Upon exposure to near UV (350–400 nm), electron-beam or X-rays, a Lewis acid is generated which induces cationic polymerization. In common with other photo-resists, SU-8 can be structured by partial exposure and the unexposed areas are subsequently removed by dissolution in a developer/solvent. The high functionality of SU-8 provides fast and extremely dense cross-linking, which suppresses diffusion of the photo-initiator from exposed to unexposed regimes. High contrasts can be achieved enabling structures to made that have at aspect ratios up to 50:1 [771,772]. The precursor has good solubility in organic solvents and, by varying solution viscosity, film thicknesses up to 700 μm can be deposited with a single coating [742]. Layers of up to 2 mm thickness can be obtained by multilayer spin-coating [773]. Limitations for the structural use of SU-8 may arise from the large thermal coefficient mismatch with most substrates, e.g. silicon. This can lead to large stresses along the layer-substrate interface after processing and delamination of the SU-8 layer from the substrate [773,774]. SU-8 is a highly brittle material in its fully cured state [775]. Fracture samples have shown almost no plastic deformation [741].
Table 11
Material properties of selected structural polymers

<table>
<thead>
<tr>
<th>Material</th>
<th>Young’s modulus (GPa)</th>
<th>Tensile strength (MPa)</th>
<th>Glass transition (°C)</th>
<th>Thermal degradation (°C)</th>
<th>Thermal conductivity (W/(m K))</th>
<th>Thermal expansion (μm/(m K))</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU-8</td>
<td>0.7–5 [741,774–777]</td>
<td>15–53 [775,833]</td>
<td>&gt;200 [740,775,778]</td>
<td>&gt;380 [740,775,778]</td>
<td>0.2 [825]</td>
<td>52 [826]</td>
</tr>
<tr>
<td>PANI</td>
<td>0.1–2 [795,797–799]</td>
<td>0.5–50 [839,840]</td>
<td>100–160 [797,807]</td>
<td>&gt;200 [841,842]</td>
<td>0.01 [795]</td>
<td>N.A.</td>
</tr>
</tbody>
</table>
The Young’s modulus of SU-8 lies within the range of 0.7–5 GPa, depending on curing conditions, which can be varied to give different cross-link densities [741,774–777]. A great advantage for high temperature applications is the inherent high thermal stability of SU-8 featuring a degradation temperature of ~380 °C and a glass transition temperature above 200 °C at maximum cross-link density [740,775,778]. Excellent chemical resistance is reported for SU-8 [741]. There are some indications that SU-8 can be classified as biocompatible material, however, to-date it is not clear if this material meets the tight requirements of certain standards for medical applications, for example, the ISO 10993 requirements for medical implants [779,780].

Photosensitive polyimide was first reported in 1971 by Kerwin and Goldrick [781]. As polyimide is generally insoluble in many solvents, it is usually prepared by imidization of photo-sensitive precursors based on poly(amic acid)s or poly-isoiimides, via thermal or chemical treatment [782,783]. Spin casting solutions of varying viscosity yields films with thicknesses up to 50 μm with a single coating [784,785]. Structuring of polyimide films is described in detail in a previous section of this review and structures with aspect ratios of 20:1 have been reported [786]. Polyimide is a ductile material which allows for maximum elongation up to 70% [782]. The Young’s modulus of photosensitive polyimide is usually within the range of 2.5–4.7 GPa for commercially available materials [787], but it can be greater than 7 GPa using optimized photosensitive groups [782]. The most noteworthy characteristics of polyimide are its excellent thermal stability and high toughness. It can be elongated up to 45% before break at 200 MPa in tension [787], and it exhibits a glass transition temperature as high as 290–350 °C in the fully cured state [787]. Thermal degradation of polyimide starts at temperatures above 420 °C [781,787]. In addition, polyimide also offers superior chemical and electrical barrier properties [788].

17.1.2. Thermoplastics

Thermoplastics refer to those polymeric materials (including the above mentioned ‘standard’ polymers) which soften or melt upon the application of heat. They are further separated into those which are (semi-) crystalline and transfer into a viscous liquid above their melting temperature, and those which are amorphous and transfer into a viscous liquid above their glass transition temperature. They can be processed at the micro-scale by soft-lithographic techniques at an appropriate temperature, for example, by embossing, molding or imprinting. The feature size of the micro-structures then replicates the feature size of the embossing or imprinting stamp, or the mould. Moulds are usually made from Si-based materials, metals, or thermosets. In this way, feature sizes in the sub-micron regime can be achieved [750,754,789]. Most thermoplastic polymers can also be structured by etching with reactive gases such as O2 or CF4 plasma [483,757]. In strict comparison with SU-8 or polyimide, thermoplastics have the general disadvantage that their long chain nature (and thus low solubility) does not allow deposition of smooth, thick films by spin casting.

Polymethylmethacryltae (PMMA) is the best characterized and most well developed thermoplastic polymer for microsystems in terms of surface micromachining [785]. However, to-date PMMA has only been used rarely as structural polymer in micro-scale sensors or actuators. Spin casting of solutions of high molecular weight PMMA results in films of thickness up to 5 μm with a single coating [790]. Thicker film deposition can be achieved by multi-layer spinning [790], or by chemical bonding of thick PMMA sheets (0.5–1 mm) onto a substrate [785]. Structuring of PMMA can be achieved by e-beam, X-ray or UV radiation, all of which cause chain scission. Afterwards a chemical development process is employed which selectively dissolves the lower molecular weight PMMA in the exposed areas. The lower molecular weight fractions dissolve much faster than the higher weight material, and variously shaped profiles can be created by careful selection and layering of differently molecular-weighted polymer. High resolution structures with feature sizes far below 1 μm can be achieved using X-ray lithography [790]. Micro-structures of PMMA can also be fabricated by soft lithographic techniques such as hot-embossing [789]. The modulus of PMMA (∼1.9–3.3 GPa [727,762,785]) is a little lower than most epoxy resins. Glass transition temperature is low, at around 105 °C and degradation temperature around 160 °C, making this a rather thermally unstable material [791]. As already pointed out, thermoplastic materials exhibit comparatively poor mechanical stability due to their visco-elastic nature. This can lead to undesirable plastic deformation of the micro-structure, particularly when the operating temperature approaches the softening temperature of the polymer.

Conjugated polymers, such as polypyrrole (PPy) and polyaniline (PANI), are widely used as structural materials in micro-actuators. The actuation of conjugated polymers is based on reversible switching between the oxidized and the reduced state, which can be accompanied by significant changes in the volume of the polymer. This process can be
driven electrochemically by small changes in cell voltage (~1 V) (please refer to the section on ionic electro-active polymers in this review). Typically, conjugated polymers are doped by ion implantation to maintain charge neutrality. The volume changes can be ascribed to mass transport of charge-balancing ions or associated solvent molecules between the polymer and the surrounding electrolyte. Owing to their low solubility, a common way of thin film deposition of PPy is monomer oxidation and the subsequent growth of polymers of the oxidized monomer [483,511,792]. Thin films also have been deposited by micro-contact-printing [793]. In contrast, PANI can be dissolved and thin films deposited by spin-casting [794,795]. Micro-patterning can be achieved by soft-lithographic methods [793] or reactive ion etching [511,792]. The modulus of conjugated polymers depends on the degree of doping, the applied potential and on the surrounding environment, for example the composition of the electrolyte solution [795–797]. Moduli within the range of 0.1–2 GPa [795,797–799] and 0.3–4.3 GPa [796,797,800–806] have been reported for PANI and PPy, respectively. The glass transition temperature of PANI lies in the range of 100–160 °C dependent on the amount of solvent in the material [797,807]. For PPy, a glass transition temperature of 42–55 °C has been reported [808,809].

17.1.3. Elastomers

Elastomers refer to cross-linked rubbery polymeric materials that can be stretched easily to very high extensions (usually 3–10 times original dimensions) and then rapidly recover to their original dimensions upon unloading. Usually, the modulus of elastomers can be tuned over a wide range by chemically controlling the amount of cross-linking between the precursor chains or monomers [729].

Among the most widely used elastomers in MEMS is polydimethylsiloxane (PDMS). This elastomer is usually synthesized by mixing a liquid monomer and a cross-linker and shaped by casting into the desired shape. Upon heating, the polymerization reaction starts and a solid elastomer is formed. Thin films of PDMS can be deposited by spin casting of the precursor liquid with thicknesses ranging from below 100 μm and 1 mm [732,810]. Structuring of PDMS can be achieved by replica molding into any desired shape [730,811]. The modulus can be tuned over a wide range (0.5–10 MPa) by varying the amount of cross-linking and the curing conditions [730–732]. PDMS is chemically inert, thermal stable, permeable to gases, low-cost and simple to handle [811]. In particular, it is an attractive material for applications in micro-fluidics, for example, as an actuated membrane, or as structural polymer in BioMEMS [732,810,811].

17.2. Applications

At a research level numerous examples of micro-scale sensors and actuators exist which benefit from the low modulus and great elasticity of structural polymers as well as from the cost-efficiency of micro-fabrication processes. However, they are restricted to only a few polymers, usually not including standard polymers. This may be due to the fact that implementation of micro-machining techniques for non-photosensitive polymers is still at an early stage of development.

17.2.1. Micro-scale sensors

The use of structural polymers for micro-scale sensors has been successfully demonstrated for chemical mass-sensing applications. These sensors make use of the low modulus and stiffness of free-standing structural polymer cantilevers. Immobilization of a monolayer of molecules on one side of the cantilever induces a difference in surface stress between the opposite sides, which results in a deflection as shown in Fig. 86. In order to promote selective binding of specific molecules, such as DNA or antibodies, the surface of the cantilever is coated in an appropriate reagent. Read-out of the sensors is usually achieved by integrated piezo-resistors [812,813]. These sensors are usually disposable and cost-efficient fabrication using structural low-cost polymers is therefore of great benefit [726]. A six times higher sensitivity can be achieved by using SU-8 as the structural polymer, compared to Si-nitride cantilevers [726]. Potential applications of thermoplastic polymers with even lower modulus and greater cost-efficiency have not yet been evaluated systematically to the current knowledge of the authors.

Limitations may arise due to electron migration in the piezo-resistive layer at high supply voltages, leading to an increase in the minimum deflection that can be detected [812]. Self-heating of the cantilever could result in plastic deformation of the structural polymer [812] if the operating temperature approaches the softening temperature.

Micro-scaled sensors cantilever arrays have been fabricated using Epo-tek UVO-114, an optically transparent epoxy, which mimic the biological front-end processing in the mammalian cochlea. They are intended for use in
auditory prostheses. Fabrication of the epoxy cantilevers was achieved by microinjection-molding using a flexible PDMS mould [814].

Another application of structural polymers for micro-scale sensing applications is the fabrication of epoxy soft tips for scanning force microscopy [815,816]. The reduced modulus of SU-8, compared to conventional Si-based tips, allows for a low-cost fabrication of soft tips with non-critical dimensions. SU-8 has also been utilized as structural material for wells suitable for large volume deposition of gas sensitive films onto a constrained area [817].

17.2.2. Micro-scale actuators

Micro-scale actuators can also make use of the low modulus of structural polymers and larger displacements at smaller actuation forces can be achieved. Actuated polymer-microstructures can be divided into two groups. The first group comprises polymer structures that can be directly actuated using some special electrical, electrochemical or thermal properties of the structural polymer. Electronically conducting polymers or conjugated polymers, such as polypyrrole (PPy) [483,495,818–820] or polyaniline (PA) [795,799,820–822], have been used as structural materials for biomedical devices. The applications are in cell manipulation [795], mimicking of muscle-like movements [792,495,823,824] or actuation of micro-valves in microfluidic systems [734]. The great advantage of these actuators, compared to the widespread piezoelectric actuator devices, lies in the low voltages that are required to drive them [820]. Underwater applications are possible since electrolysis due to high voltages is no longer a concern [795]. The excellent thermal properties of highly cross-linked epoxies, such as SU-8, can be used for thermal actuation [825]. Due to their large thermal expansion coefficient, SU-8 structures can be actuated at low voltages <1–2 V and at low ambient temperatures (10–32 °C), which allows for actuation even in physiological solutions [826]. Actuators incorporating piezoelectric polyvinylidene fluoride (PVDF) have also been reported [827].

The second, larger, group of actuated polymer micro-structures comprises the majority of structural polymers. These exhibit rather poor electrical, electrochemical, or thermal properties and thus they have to be actuated indirectly. For example, metal coated polymer structures made from polyimide or SU-8, can be actuated electrostatically by using the metallic layer as an electrode [828,829]. Pneumatic actuation of SU-8 micro-grippers has been reported which benefit from the low modulus and the good transparency of SU-8 for visually controlled micro-assembly applications [724,830], Fig. 87. Low-modulus (0.75 MPa) elastomeric membranes of polydimethylsiloxane (PDMS) with attached magnetic Permalloy (Ni80Fe20) films can be actuated magnetically, enabling large displacements for applications of tetherless micro-pumps in microfluidic systems [810].

Although numerous instances of structural polymers used in micro-scale actuators can be found in literature, the general benefits that polymer materials offer through their exceptional variety of properties have not been exploited to full capacity. A novel actuation mechanism is now under investigation which permits direct actuation of non-conductive structural polymers [831]. This uses the so-called Kelvin polarization force, which acts on a dielectric structural polymer in the presence of an electric field gradient. A new fabrication process for polymer micro-cantilevers (Fig. 88) and micro-bridges has been developed, which is based on conventional micro-machining
techniques. Using a combination of photolithography and sacrificial layer etching, similar to the approach for structuring of PMMA-based micro-bridges [757], this fabrication process can be applied to virtually any polymer. Finally, a new micro-scale, wireless, passive strain sensor for biomedical applications has been proposed, which can be built entirely from low-cost thermoplastic polymer materials [832].

18. Integration and interconnection

Integration of transducer materials with their associated electrical and structural components is an essential consideration for any micro-sensor or actuator device. It has been roughly estimated that packaging alone accounts for between 50\% and 90\% of the cost of a commercial device and it is responsible for between 50\% and 75\% of failures. In reality, whilst some standard device packages are offered by commercial foundries, it must be said these are not

\[\text{Cristina Rusu, Katrin Persson}\]
Commercially there are benefits to be had from adopting an integrated approach to MEMS design, where all aspects of a complete system are considered in the conceptual design phase. In practice, this type of approach is rarely adopted because of its inevitable complexity. More commonly, devices are created as a series of individual component parts for later integration. On the basis that one man’s system is another man’s component, it is not, therefore, possible in this review to discuss the vast range of materials that may find usefulness in wider aspects of MEMS device technology. In this final section, it is proposed to focus on two aspects of fabrication that have important consequences for the creation of novel devices and which are widely applicable for integration of new transducer materials into useful systems. The first of these relates to wafer bonding, which is an essential tool for the creation of hybrid multi-material structures. This is presented in overview. The second part of this section discusses newly-available materials for low thermal expansion substrates in the form of low temperature co-fired ceramics (LTCC). These provide excellent capability to create vias and/or channels and LTCC wafers can be stacked to build up structures with 3-D functionality.

18.1. Wafer bonding

Most micro-fabricated structures are developed at the die or wafer scale and assembled by bonding. In this way, it is possible to build up new three-dimensional structures and complicated devices that incorporate several different materials, examples being glass on silicon, polymers on Si/SiO₂, GaAs on Si, InP on Si. Individual components can be prepared separately and procedures can be performed either before or after bonding depending on the process requirements. Bonding can be also used for packaging and encapsulation on the wafer scale (this is termed ‘zero-level packaging’) and at die-level, for instance by flip-chip bonding. In general, wafer level bonding allows for higher process flexibility than die-level bonding as the local environment can be more fully controlled, enabling gas-filled cavities or vacuum chambers to be formed. By extension, ‘first level packaging’ refers to chip encapsulation and the leads used to interconnect the chip to the outside world. These terms are derived from earlier IC technology. In contrast to IC devices, zero-level packaging is necessary to protect many micro-mechanical devices at the wafer scale before they are diced into individual chips. This is particularly true when the devices include free-standing structures or moving parts, examples being inertial systems (accelerometer, gyroscopes), bolometers, mirrors and switches. All of these MEMS devices can benefit from hermetic sealing and/or enclosure in a controllable atmosphere to ensure good operation and lifetime [844].

The most common bonding methods used in microtechnology are adhesive bonding, thermo-compression bonding, eutectic bonding, glass-frit bonding, solder bonding, silicon direct bonding (or fusion bonding), plasma-enhanced bonding and anodic bonding (or electric-field-assisted bonding). This wide variety of different techniques has emerged owing to the multitude of applications and the different bonding requirements that can be encountered [845,846,857].

18.1.1. Adhesive bonding

Adhesive bonding with polymers allows a free choice of wafer materials with different coefficients of thermal expansion (CTE). Common adhesives that can be used are: (i) thermosetting polymers (curing based on chemical reaction), such as BCB (benzycyclobutene <350 °C), SU8 (<230 °C), polyimide (<350 °C), negative resist (<200 °C) and (ii) thermoplastic polymers (processed by heating and cooling), such as PMMA (polymethyl methacrylate, easily etchable) and liquid crystal polymers (low moisture uptake). On the wafer scale, it is very difficult to remove incipient gases and the adhesive material should ideally generate no by-products on curing. BCB achieves this by a ring-chain polymerisation mechanism. The adhesive materials are deposited mainly by spinning (alternatively by lamination, evaporation, screen printing or spraying) to obtain uniform coatings between 0.1 and 50 μm. Patterning by photolithographic methods (using a photoresist mask and dry etching) can be used under suitable conditions can be performed to bond in selected areas and post-bond annealing at low temperature (below 350 °C) is sometimes used to reduce stresses in the bonded stacks. Satisfactory bond strengths up to 10 MPa can be obtained [847–849].

Polymer bonding does not produce a hermetic seal [850], and in practice polymers tend to exert a high vapour pressure that will reduce an encapsulated vacuum. For hermetic sealing, a metallic bonding process is normally required. However, they are very versatile and ideal for a vast range of applications, for example, in micro-fluidic devices. They can be biocompatible and chemically resistant and importantly they can be applied at relatively low temperatures. They can adjust more easily than other techniques to variations in wafer topography, selective bonding
of individual components on wafer areas is possible, and manufactory costs are low. On the negative side they can exhibit reduced alignment accuracy due to slippage or differential thermal expansion when thick layers are used.

A further application area for polymers is for temporary adhesion bonding, used as an intermediate step in processing. This is particularly useful for grinding down and polishing of the wafers, for transfer bonding [849] and when ultra-thin devices are fabricated [851].

18.1.2. Metallic bonding

Metallic bonding includes solder, eutectic, and thermo-compression bonding (Table 12). InSn and SnPb [852,853] are common solder materials used for single-chip devices and bond strengths up to 20 MPa have been obtained [854,855].

The Si–Au system is usually considered the most promising for eutectic bonding, because of its low eutectic temperature, its widespread use in die bonding and its compatibility with aluminium interconnections [856]. A disadvantage of Au–Si eutectic bonding is the difficulty of obtaining complete uniform bonding over large areas, and native oxides are sufficient to prevent bonding from taking place [856]. Eutectic bonding is reported to introduce substantial mounting stress in piezoresistive sensors, causing long-term drift due to relaxation of the residual stresses [856]. Localized Au/Si eutectic bonding has been successfully achieved at temperatures above 800 °C by applying 1 MPa contact pressure [858]. The local heating was provided by specially constructed micro-heaters.

Au–Au thermo-compression bonding takes place at temperatures between 300 and 400 °C and pressures of 10–100 MPa giving very strong bonds [859]. Micro-mirrors were bonded to a CMOS chip at 300 °C applying 100 MPa pressure [860].

Metallic bonding has the advantage of easily forming a hermetic seal at moderate process temperatures (200–400 °C), but the processes are intolerant of variations in wafer topography and making electrical feed-throughs can be complicated.

18.1.3. Glass-frit bonding

Glass-frit bonding also uses the thermo-compression method. Glass-frit is obtained in the form of a paste containing glass beads. It is deposited by screen-printing onto a substrate and solvents are removed by degassing. This is followed by the application of heat and pressure to make a hermetic seal. Glass-frits are available that are thermally matched to silicon. The compression ratio is ~3:2 after bonding.

The advantages of glass-frit bonding are high flexibility, economical processing, free choice of wafer material, few constraints on the bond surfaces (it adapts itself to the wafer topography). However, glass frit printing can be a messy process.

18.1.4. Silicon direct bonding

Fabrication of silicon-on-insulator (SOI) wafers was initially the main motivation for studying silicon direct bonding (SDB) [861]. Manufacture of SOI wafers with buried cavities makes it possible to take advantage of the mechanical and electronic properties of single-crystal material in the relatively thin device layer [862].

SDB is also called silicon fusion bonding but this name is misleading because the (annealing) process temperatures, 800–1100 °C, are below silicon’s fusion point, 1410 °C.

Two silicon surfaces, silicon and silicon oxide layers or thin layers of silicon nitride (100–200 nm), two silicon oxide layers and surfaces of silicon nitride are typical surfaces in between which SDB is possible. It is essential that the bonding surfaces are very clean. This can be achieved using RCA1 or Piranha etch (H₂SO₄:H₂O₂) to remove organic contaminants and RCA2 to remove metal and alkali contaminants. The surfaces then tend to become hydrophilic in character with a native oxide and a high density of OH groups. By dipping in hydrofluoric acid (HF) after the RCA
clean, the wafer can be made hydrophobic. Surfaces free from oxide are necessary for bonding epitaxial or hetero-
epitaxial-like silicon. The clean wafers are brought into contact at room temperature and a spontaneous pre-bonding
occurs, after which the pair is annealed at 800–1100 °C (in a wet or dry oxidation furnace). A more detailed description
of the SDB mechanism can be found in [846,863]. For hydrophobic surfaces, which are terminated mainly by
hydrogen and some fluorine after the HF dip, the bonding mechanism is not yet understood [864,865]. The bond
strength it can be as high as the fracture surface energy of bulk silicon (2.5 J/m²) at annealing temperatures above
900 °C for hydrophilic surfaces. For hydrophobic wafers, the bonding energy is very low (13–26 mJ/m²) for the
spontaneous bonding (no anneal).

Examples of different material combinations fabricated by direct wafer bonding include silicon on sapphire (SOS),
SrTiO₃ and LiNbO₃ for optical applications [866], crystalline quartz on silicon for high frequency applications, Si on
fused quartz for high-definition television (HDTV) projection masks [867], GaAs or InP on silicon for combining
opto- and microelectronics, different III–V compounds for fabrication of vertical cavity surface emitting lasers
(VCSEL) and light emitting diodes (LED) [868]. The Smart-Cut process [869] is used for metallic direct bonding
to transfer Si, GaAs and InP onto a silicon substrate but it applies to other materials [870], SiC, InP, GaAs, Ge,
diamond.

18.1.5. Plasma-enhanced bonding

Materials with low decomposition temperatures, materials with different thermal expansion coefficients or other
temperature sensitivity (even if they are otherwise thermally matched) can be joined by low temperature bonding.
Sensitive mechanical and electrical components are often integrated on the same chip to shorten the distance from the
mechanical output to the detection and analysis electronics and to minimize losses due to leakage and parasitic
capacitances.

For low or room temperature wafer bonding, plasma pre-treatments of wafers have recently attracted attention. The
bond strength of direct bonded wafers at room temperature has been shown to increase after plasma treatment.

A variety of plasma systems can be used. The most common plasma, oxygen, is known for effective removal of
hydrocarbons and water related species from wafer surfaces [871,872]. However, if applications require oxide-free
interfaces fluorine-containing plasmas [873], SF₆ containing plasmas [874] or Ar containing plasmas [875] can also be
used.

The results of the different investigators cannot be compared directly because experiments have been made with
such a wide range of plasma systems and different wet chemical treatments. The physical mechanism that improves
the surface energy of plasma treated bonded wafer pairs compared to wet chemical bonded ones is still under debate
[873,876–878].

A wide variety of different materials can be bonded to each other using low temperature plasma enhanced wafer
bonding, such as Si, SiO₂, glass, sapphire [879], quartz [880] and InP [873,898].

Beside measurement of surface energy and bond strength [881,882], gas leakage measurements have been made on
bonded wafers indicating that strong and hermetically sealed bonds can be obtained by applying plasma enhanced
bonding [883]. However, the impact on the electrical properties of the bonded interface and on the reliability of
standard electronics due to plasma treatment is unclear [884,885].

18.1.6. Anodic bonding

Anodic bonding is performed using an electric field (200–1000 V) and raised temperature (180–450 °C), but with
no mechanical force. It can be performed in vacuum or at atmospheric pressure. Typical materials include slightly
conductive glass wafers and virtually any metal [886] or Si wafer. Otherwise two Si wafers with an intermediate thin
film glass [887] can be bonded anodically. The glass used for anodic bonding contains a high concentration of alkali
metal oxides and has a similar CTE (ca. 3.2–3.6) to Si (2.3–2.6). The most used glass types are borosilicate (Pyrex/
Corning7740, Schott 8330, Schott borofloat33), lithia potash borosilicate (Corning7070 this has a less favourable CTE
match to Si, but it is useful when thick oxide/nitride has to be bonded), and Hoya SD-1, SD-2 (this has a better CTE
match to Si, leading usually to less residual stress, but it needs higher temperature/voltage to achieve high bond
strength). At moderate temperature, the glass becomes a conductive solid electrolyte and due to the applied electric
field the positive metal ions become mobile and migrate towards the negative electrode. The high electrical field
between the glass and the silicon creates an electrostatic force pulling the wafers into contact resulting in an oxidation
procedure to form covalent bonds. Electrochemical, electrostatic and thermal mechanisms and combinations of these
have been suggested to explain the bond formation [888]. The bond strength of 10–15 MPa up to 30–40 MPa has been obtained [889] and it was shown that it increases with time and temperature [890].

Anodic bonding is mostly used in wafer-level packaging and device encapsulation [891–893] due to its hermetic seal, high mechanical and chemical stability. However, it is not easy to achieve a high vacuum in the cavities. Two potential sources of gases are the inner surface of the sealed cavity and oxygen release (O₂) during the bonding process [894].

A thin layer of sputtered glass [894,895], electron beam evaporated glass or spin-on glass can be used instead of a glass wafer. The advantages of this approach are a decrease in residual stress (low difference in CTE between the bonded wafers) and in bonding voltage. Strong bonds have been achieved with voltages as low as 20–40 V [894]. The bond strength was between 2 and 3 MPa (measured by pulling the wafers apart). Successful bonding was also obtained when the second Si wafer was covered by a thin layer of oxide (100 nm), nitride (100 nm), polysilicon (500 nm) and aluminium (0.4 μm) [896].

The advantages of anodic bonding are its reasonably low process temperature (below the sintering temperature of standard Al–Si metallisation), low residual stress, hermetic seal and less stringent requirements on the surface quality of the wafers (as compared to silicon direct bonding). The surface roughness typically must be smaller than 1 μm rms. The surface should also be clean and dust free. Despite low residual stresses and relatively low process temperatures anodic bonding is not suitable for all applications, for example where organic coatings are involved for bio-applications. Moreover, the high electrical field and sodium contamination are disruptive for MOS devices [897].

### 18.2. Low temperature co-fired ceramics and microsystems

Almost all microsystems obtained by wafer/chip bonding need to have electrical interconnections and these are often difficult to achieve. Buried feed-throghs are mainly formed by surface micromachining and this requires extra processing steps. Different bulk micromachining methods are also used in combination with Si and/or glass bonding (Fig. 89). One method of improving the electrical connections and at the same time making the fabrication process easier is to use LTCC (low temperature co-fired ceramic) material for 3-D interconnections and to replace glass in hermetic packaging.

The combination of ceramics with micro-scale components opens up new possibilities for emerging applications and new products by exploiting the ability of 3-D ceramic structures to integrate and interconnect with packaging and with microfluidic, optical, micro-reactor and sensing functions. Low temperature co-fired ceramic (LTCC) can take the form of multilayer structures or thick film hybrids. Combined technologies such as selective transfer bonding and laser machining can also be used to create new geometries and embedded passive components. The benefits are increased circuit density, enhanced functionality, improved performance and low-cost prototyping. Emphasis is

![Cross sectional sketch of a triple stack (left) glass-Si-glass with squeezed metal electrode, (right) SOI with lateral electrical contact.](image-url)
currently being placed on material studies and processing along with prototype development, advanced design and application opportunities [899]. The ceramics can be highly cost effective and reliable, providing high performance interconnections in products that operate in hostile thermal and chemical environments. Wide ranging applications therefore exist in the automotive and aerospace industries and also for MEMS, sensors, wireless and optical communications.

Recent instances of LTCC used in hybrid combination with microsystems at the meso-scale are described in references [900–903]. A microfluidic application, wherein part of the MEMS function is integrated into the LTCC carrier (gas and liquid channels) is illustrated schematically in Fig. 90. The LTCC offers full electrical interconnection capability to any point on the sensor, and it provides excellent mechanical stability.

18.2.1. Medium CTE LTCC

The standard LTCC material is green tape LTCC 951 from Du Pont which has a coefficient of thermal expansion (CTE) of 5.6 ppm/K, it contains 50% alumina and the remainder is mainly borosilicate glass. Shrinkage is about 12.7% in the x, y plane and 15% in the z-axis. This material does not allow for direct bonding to Si (mainly due to the large mismatch in CTE), but by using an intermediate layer of glass as a ‘buffer’ layer it may be possible to combine it with Si. The glass paste from Corning 7070 (developed by DuPont) is suitable for anodic bonding and it can be selectively deposited on the substrate in any pattern by either screen or stencil printing. These printing techniques are well established, simple, and low-cost, and they are utilized in many applications, including packaging.

Tests have been performed by the authors firstly bonding Si wafers, and bonding Si to LTCC where one wafer had a screen-printed glass grid patterned onto it. Screen-printing and stencil printing were successful on both the silicon and LTCC substrates (Fig. 91). However, the waviness of the LTCC substrates proved much more significant in comparison to silicon, thus screen-printing was more difficult. The glass was fired at 900 °C for 60 min and slowly cooled down. The glass line widths were about 1 mm, and thicknesses were 10–35 and 35–100 μm for the screen-printed glass and stencil-printed, respectively. The glass surface roughness was about 1 μm. For good anodic bonding, the glass surface needs to be polished. The thickness and profile of the fired glass depended on the chosen pattern, printing method (screen or stencil), firing temperatures and durations. These factors in combination with the properties of the glass paste can be modified to create an even smoother surface.

Two test samples, Si-glass paste-Si and LTCC-glass paste-Si, were bonded anodically at 420 °C using 130–500 V. The scanning electron microscopy (SEM) cross-section of Si-glass paste-Si showed that almost 50% of the width of the original printed glass had resulted in a successful bond. For the LTCC-glass paste-Si, fracture seems to occur at the Si-glass interface during cooling after bonding (Fig. 92) [904]. Low-temperature anodic bonding has previously been reported [905,906]. The key issue to address is successful matching of the coefficients of thermal expansion and hermetic sealing can be accomplished through adjustments to the composition of the glass [907].

Fig. 91. Screen-printed glass checked pattern 6 × 6 mm grid, 1 mm wide lines: (a) silicon wafer and (b) LTCC.
Cavities within LTCC can be created simply by cutting the LTCC tape. No sacrificial material in the cavities need be used during lamination and subsequent sintering (Fig. 93). The large cavities (>3–4 mm²) do not collapse during the normal LTCC manufacturing procedure but some distortion can occur. Si pieces have been encapsulated into the LTCC layers without breaking and the silicon becomes mechanically clamped in place when the ceramic tape is fired. Fluidic connection is another area where LTCC can be used. Alumina tubes (800 μm outer diameter) were successfully connected to LTCC to form fluidic channels (Fig. 94). Again the shrinkage of LTCC can be exploited to provide mechanical fixing and different tube sizes can be fitted.

18.2.2. Low CTE LTCC

HITK and VIA Electronic (Germany) are developing a new LTCC material that has a lower CTE, 3.6 ppm/K. The new LTCC material contains alkali ions that permit direct anodic bonding to silicon [903]. This should make it possible to obtain good packaging, with controlled atmosphere inside cavities and easy access for electrical connections, through anodic bonding of LTCC to Si (Fig. 95). The LTCC can stand high temperatures (few hundred degrees) and it is chemically inert. CTE is slightly higher than glass, at ~3.6 (Fig. 96) and therefore it is unsuitable for high performance devices such as those used for inertial applications due its excessive thermal mismatch.

---

Fig. 92. Fractures during cooling occurring in Si: (a) screen-printed glass paste on LTCC and (b) silicon wafer after fracture.

Fig. 93. IR image of empty cavities embedded in LTCC.
The low-CTE LTCC can have the alkali ions in the bulk or only at the surface. Trials were performed by the authors using bulk alkaline material. A 2-in. wafer was bonded to silicon by standard anodic bonding at 420 °C, 800 V in less than a minute (Fig. 97) [908] and inspected with infra-red light. A void can be seen in this example, but this is due to unwanted scratches on the LTCC wafer.

Metal electrodes embedded in an LTCC substrate (placed below the surface) permitting lower bonding voltages and/or confined electric fields to be used for applications containing more critical structures (Fig. 98). An LTCC substrate was successfully bonded anodically to Si wafer at 420 °C and 30 V within a few seconds.

18.3. Characterisation methods for microsystem bonding

There are several ways to characterise these types of bonds: (i) non-contact techniques (infra-red, ultrasonic imaging, X-ray diffraction, topography, scanning acoustic microscopy); (ii) fracture surface energy testing (Maszara test); (iii) bond strength testing (pull test, burst test); (iv) hermeticity (membrane test, helium-leakage, resonator/pressure).

Infra-red (IR) light can be used to detect voids (non-bonded areas) larger than 20 μm (sample thickness must be larger than half the wavelength of the light and void length larger than the numerical aperture of the optical system).

![Comparison of the coefficients of thermal expansion for LTCC BGK 79 (VIA Electronic), Si and Borofloat 33.](image)
Infra-red measurements, such as multiple internal reflection geometry-Fourier IR spectroscopy (MIRS-FTIR) can be used to investigate the chemistry of the bonded interface [909]. Ultrasound imaging (scanning acoustic microscopy) can give further qualitative information about bond quality. X-ray diffraction, topography and scanning acoustic microscopy offer an increased lateral resolution of 1 and 4 μm, respectively [910].

Fracture surface energy can be measured using a method introduced by Maszara [911]. A thin blade is inserted between the bonded wafers and a crack at the interface is introduced. The length of the crack provides a measure of the surface energy $W$ of the bond:

$$W = \frac{Et^3y^2}{8L^4}$$

where $E$ is Young’s modulus of the wafer material, $t$ the thickness of the wafers (equal thickness of the wafers assumed), $2y$ the thickness of the blade and $L$ is the crack length. This equation is particularly valuable when a pair of similar wafers are bonded. The Maszara method has the advantage of creating a well-defined load on the bonded interface. Unfortunately, uncertainties in the crack length result in large errors (surface energy is a fourth power of the crack length).

Bond strength can be measured by pull tests [912,913]. Diced samples are glued onto stud poles that are pulled apart in a controlled manner. However, sample mounting is critical and the results depend strongly on it. A better method can sometimes be burst tests [914].
Hermeticity tests of electronic devices are usually carried out using helium leak detection [850,887] or using pressure sensor/resonators inside the cavity (quality factor of a resonator depends on the pressure) [915]. Jourdain found no fine or gross leak for BCB bonds when tested according to MIL-STD-883D method but the performance of MEMS devices inside the cavity suggested a leak. It was concluded that new standards should be created for MEMS devices to account for their small volumes [847].

For the characterisation of the electrical properties of bonded interface, standard electrical characterisation methods such as IV-, CV- and resistance measurements are commonly used [916].

18.4. Conclusion

Wafer bonding technology allows parallel fabrication of multiple chips and extends freedom of design for 3-D devices to a multitude of materials. The choice of bonding method depends both on the application and on the need for compatibility with subsequent stages of processing. Bonds can be achieved that can withstand harsh environments, for instance, high temperatures or high humidity, implantation in the human body or in the combustion motor of a rocket.

Low temperature co-fired ceramics (LTCC) are proving to be exceptionally useful. They can be used as substrates, fluidic circuit boards, for electrical inter-connectors and in packaging. New low-CTE LTCC substrates permit hybrid integration through direct bonding to MEMS wafers thereby providing full 3-D functionality. They will permit cheaper chip and wafer scale packages to be made and they can therefore play a major role in bringing MEMS products to a mass-market due to lower production costs.

Acknowledgements

This work was carried out within the framework of the EC Network of Excellence ‘Multi-Material Micro Manufacture: Technologies and Applications (4M)’. The authors gratefully acknowledge the support provided by the 4M Network in the preparation of this article. Particuluar thanks go to the leader of the Sensors and Actuators Division, Dr. Paul Kirby and to Enza Giaracuni in the Microsystems and Nanotechnology Centre at Cranfield University for her work in preparing the manuscript. Stephen Wilson would like acknowledge work carried out under the EC Growth Project ‘Aeromems II’ in collaboration with BAE Systems Advanced Technology Centre, Sowerby, Bristol, UK and the financial support of the UK Engineering and Physical Sciences Research Council for the project ‘Piezoelectric Micro-actuators for Smart Micro-Air Vehicles’. In addition, the support of DSTL and EPSRC under the Joint Grants Scheme is acknowledged for the project ‘Electric-Field Structuring of Piezoelectric Array Composites’. Emmanouel Pagounis would like to thank Dr. Herbert Schmidt from Robert Bosch GmbH for providing Fig. 22 in the text. Michael Wegener is indebted to Professor Dr. Reimund Gerhard-Multhaupt, Dr. Axel Mellinger, Werner Wirges, Dr. Guggi Kofod, Andreas Pucher (Potsdam), Dr. Enis Tuncer (formerly Potsdam), Prof. Dr. Siegfried Bauer, Dr. Simona Bauer-Gogonea, Dr. Reinhard Schwödauer (Linz), Prof. Dr. G.M. Sessler, Dr. R. Kressmann, Dr. J. Hillenbrand (Darmstadt), Dr. M. Paajanen, Dr. J. Raukola (Tampere), Prof. Dr. Gian Carlo Montanari, Dr. Davide Fabiani (Bologna), Prof. Dr. Zhongfu Xia and Dr. Xunlin Qiu (Shanghai), for exciting and fruitful discussions on the topics of this paper and the excellent co-operation in joint research projects. Partial financial support from the German Science Foundation, the Volkswagen Foundation, the German Academic Exchange Service, the German Federal Ministry of Economy and Labor, and the European Commission is gratefully acknowledged.

References