

RWTH edition



**RWTHAACHEN**  
UNIVERSITY

Werner Karl Schomburg

# Introduction to Microsystem Design



Springer

RWTHedition

RWTH Aachen University



Werner Karl Schomburg

# Introduction to Microsystem Design

 Springer



Prof. Dr. rer. nat. Werner Karl Schomburg  
RWTH Aachen University  
Konstruktion und Entwicklung  
von Mikrosystemen (KEmikro)  
Steinbachstraße 53 B  
52074 Aachen  
Germany  
schomburg@kemikro.rwth-aachen.de

ISSN 1865-0899                      e-ISSN 1865-0902  
ISBN 978-3-642-19488-7            e-ISBN 978-3-642-19489-4  
DOI 10.1007/978-3-642-19489-4  
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011928914

© Springer-Verlag Berlin Heidelberg 2011

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Cover design:* deblik, Berlin, Germany

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

Microsystems technology (MST) or Microelectromechanical systems (MEMS) how it is called in America is a comparatively young emerging technology, which allows building up miniaturized devices such as microvalves for implantable medicament dosing systems or micro-total analysis systems ( $\mu$ TAS) which shall provide a miniaturized laboratory on a polymer chip just a few centimeters in size. The first steps to MST had been done more than 30 years ago when anisotropic etching of silicon was discovered [1] and sacrificial layer technique was invented [2].

Nowadays, MST is a well-established technology which is the basis of many products. Modern life in many fields is based on a variety of microsystems unnoticed by most of us. In most cars, microsensors for the measurement of acceleration, yaw rate, pressure, and flow are implemented. Watches, hearing aids, mobile phones, beamers, ink-jet printers, PCs, and catheters for minimal invasive surgery are other examples of applications which became possible in the present form by microtechnologies only. Accordingly, many jobs are available in microtechnique and much more jobs are depending on it.

In previous decades, the fabrication techniques of MST had been the main issue of research and development resulting in today's more or less standard production processes such as bulk silicon etching, reactive ion etching, surface micromachining, micromolding, silicon fusion bonding, etc. These processes are well described in several text books [3–6], and, therefore, are available for both industry and teaching at universities.

However, MST is not only characterized by its novel fabrication processes. The transition to smaller dimensions is combined with the need for a change in design also. A miniaturized sensor or actuator requires a different design due to both the new fabrication techniques and the smaller scale which results in a change of the significance of effects and forces. For example, capillary force is of no importance in the macroscopic world, while it may be used as the driving force in microscopic designs. The piezoelectric effect and thermal strain are known and need to be considered in macroscopic engineering but play a much more important role in MST.

Until now, there is no textbook which describes the design of micro systems systematically. Therefore, this book was written to fill this gap. It is based on a course given at RWTH Aachen University and Tsinghua University in Beijing for undergraduate students in their fifth or higher semester. This book may be used as

the basis for similar courses, for self-study, or as a reference for the experienced engineer. All the equations presented here are not limited to microsystems but are valid in general. Therefore, this book may help also engineers working in different fields.

This book does not describe the fabrication processes of MST but can be understood without knowing these processes. It provides the basic equations needed to calculate or at least estimate the order of magnitude of the effects and forces which are important in MST. For quick reference, these equations are presented in tables which are found in an index on page xxi.

# Notations and Symbols

The following notations and symbols are used throughout this book. The notations and symbols are not always defined in the near of an equation avoiding interruption of the text. If the meaning of a parameter is not obvious to the reader, it will be found in this list.

$A$	Area
$A_B$	Cross-sectional area of a beam
$A_C$	Inner area of a capacitor electrode
$A_F$	Cross-sectional area of a fluid
$A_g$	Cross-sectional area of a gap
$A_i$	Amplitude of mode $i$
$A_K$	Area of a body
$A_M$	Area of a membrane
$A_{n,m}$	Amplitude of the mode $n,m$ of a membrane
$A_P$	Area of a piezo
$A_S$	Area of a layer
$A_Z$	Cross-sectional area of supply and delivery pipes
$A_0$	Amplitude
$a$	Coefficient
$a_C$	Coriolis acceleration
$a_F$	Acceleration of a fluid
$a_i$	Coefficients
$a_M$	Length of a rectangular membrane
$a_m$	Acceleration of a mass
$a_p$	Coefficient
$B$	Width of a slide valve
$B_{F,opt}$	Width of a slide valve optimum with respect to small switching force
$B_{W,opt}$	Width of a slide valve optimum with respect to small switching energy
$b_B$	Width of a beam
$b_C$	Width of a capacitor electrode
$b_K$	Width of a capillary
$b_L$	Width of a conductor path

$b_M$	Width of a rectangular membrane
$b_{PZ}$	Width of a Pockels cell
$b_P$	Width of a piezo
$b_p$	Coefficient
$b_S$	Width of a layer
$B_Z$	Width of supply and delivery pipes
$C_D$	Heat capacity of a wire
$C_{el}$	Electrical capacity
$C_{fl}$	Fluidic capacity
$C_P$	Heat capacity of a gas at constant pressure
$c$	Speed of sound or light
$C_{th}$	Heat capacity
$C_V$	Heat capacity of a gas at constant volume
$c_F$	Concentration of particles in a liquid
$c_f$	Specific heat capacity of a fluid
$c_p$	Specific heat capacity of a gas at constant pressure
$c_v$	Specific heat capacity of a gas at constant volume
$D$	Diffusion constant
$D_i$	Charge density in $i$ -direction
$D_h$	Hydraulic diameter
$D_{h,Z}$	Hydraulic diameter of supply or delivery pipes, or feed channels
$D_L$	Diameter of a lens
$D_R$	Damping constant
$D_S$	Diameter of a valve seat
$d$	Distance or diffusion depth
$d_B$	Thickness of a beam
$d_C$	Distance of capacitor plates
$d_E$	Thickness of electrodes
$d_f$	Thickness of a thin film
$d_{ij}$	Piezoelectric modulus of straining in $j$ -direction and loading in $i$ -direction = $\epsilon_0 \epsilon_r g_{ij}$
$d_I$	Thickness of an insulation layer
$d_i$	Distance of the electrodes of a closed relay
$d_K$	Height of a rectangular capillary
$d_L$	Thickness of a conductor path
$d_M$	Thickness of a membrane
$d_{Su}$	Thickness of a substrate
$d_0$	Distance of capacitor plates when no voltage is applied
$E_B$	Young's modulus of a beam
$E_{el}$	Electrical field
$E_f$	Young's modulus of a thin film
$E_{kin}$	Kinetic energy
$E_{kin,m}$	Maximum of kinetic energy
$E_M$	Young's modulus of a membrane

$E_P$	Young's modulus of a piezo
$E_S$	Young's modulus of a layer
$E_{Su}$	Young's modulus of a substrate
$E_T$	Young's modulus of a carrier layer
$E_V$	Ratio $(h_T E_T)/(h_P E_P)$
$e$	Elementary charge = $1.6 \times 10^{-19}$ C
$F$	Force
$F_a$	Loading force
$F_B$	Force generated by a beam
$F_{C,n}$	Normal capacitive force
$F_{C,l}$	Lateral capacitive force
$F_{el}$	Electrostatic force
$F_F$	Force acting at a frame
$F_{F,z}$	z-Component of a force acting at a frame
$F_H$	Force required for valve switching
$F_{H,max}$	Maximum force required by a valve actuator
$F_i$	Force required for acceleration or force in i-direction
$F_K$	Critical force of a beam
$F_{Ku}$	Weight force of a sphere
$F_k$	Force of a spring
$F_l$	Force longitudinal to a beam
$F_M$	Force acting on or elastic force of a membrane
$F_P$	Force generated by a piezo
$F_p$	Force generated by a pressure difference
$F_{p,z}$	z-Component of a force generated by a pressure difference
$F_R$	Friction force
$F_{sf}$	Force due to squeeze-film effect
$F_t$	Force transversal to a beam
$F_{th}$	Force of a thermal actuator
$F_U$	Force needed to flap over a beam or membrane
$F_x$	Force or force component in x-direction
$F_z$	Force or force component in z-direction
$F_0$	Force of an actuator hindered from deflection
$F_{0,B}$	Force of a bimorph hindered from deflection
$f$	Frequency
$f_A$	Frequency of an actuator
$f_B$	Focus length
$f_{F_0}$	Correction factor for the effect of the carrier layer on the maximum force of a piezo
$f_i$	Frequency of vibration mode i
$f_{n,m}$	Frequency of the mode n,m
$f_P$	Pump frequency
$f_{P,max}$	Maximum pump frequency

$f_{v_E}$	Correction factor for the effect of the carrier layer on the deflection of a piezo generated by a voltage
$f_{v_F}$	Correction factor for the effect of the carrier layer on the deflection of a piezo generated by an outer force
$f_0$	Frequency without damping
$f_1$	Frequency of the fundamental mode
$G_B$	Shear modulus of a beam
$g$	Grid constant
$g_e$	Acceleration of gravity
$g_{ij}$	Piezoelectric modulus = $d_{ij}/(\epsilon_0 \epsilon_r)$
$H$	Travel of a valve body
$H_{\max}$	Maximum stroke of a slide valve
$H_Z$	Height of supply and delivery pipes
$h_A$	Height of an actuator chamber
$h_i$	Thickness of layer $i$
$h_P$	Height of a piezo
$h_T$	Height of a carrier layer
$h_V$	Ratio $h_T/h_P$
$h_S$	Height of a layer
$I$	Area moment of inertia
$I_a$	Intensity of light at the output
$I_e$	Intensity of light at the input
$I_{el}$	Electrical current
$I_L$	Intensity of light
$I_m$	Mass moment of inertia
$I_t$	Torsional constant
$I_0$	Initial intensity of light
$i$	Index or integer number
$K$	Gauge factor of strain gauges
$K_i$	Gauge factor of strain gages in $i$ -direction
$k$	Spring constant
$k_B$	Boltzmann constant = $1.4 \times 10^{-23}$ J/K
$k_{el}$	Electrical spring constant
$k_i$	Spring constant in $i$ -direction
$k_m$	Mechanical spring constant
$k_S$	Constant
$k_W$	Heat transition coefficient
$L$	Length
$L_B$	Length of a beam
$L_{Bc}$	Length of the beam connecting two beams
$L_C$	Length of the overlap of the electrodes of a capacitor
$L_E$	Distance between electrodes
$L_{el}$	Inductivity
$L_F$	Length of a fluid column

$L_f$	Length of a thin film
$L_{fl}$	Fluidic inductance
$L_K$	Length of a capillary
$L_k$	Critical distance
$L_L$	Length of a conductor path
$L_m$	Mechanical analogon to inductance
$L_{Pa}$	Length of a parabola arc
$L_P$	Length of a piezo
$L_{PZ}$	Length of a Pockels cell
$L_R$	Distance to the rim
$L_S$	Length of a layer
$L_s$	Length of a sensor
$L_V$	Length of a valve gap in flow direction
$L_Z$	Length of supply and delivery pipes
$M$	Bending moment
$M_t$	Torque
$m$	Natural number or mode
$m_B$	Mass of a beam
$m_F$	Mass of a fluid
$m_K$	Mass of a body
$m_{mol}$	Mass of a mol
$m_T$	Mass of a particle
$m_0$	Mass attached to a beam
$m_1$	Mass of body 1
$m_2$	Mass of body 2
$NA$	Numerical aperture
$Nu$	Nusselt number
$N_A$	Avogadro constant = $6.02 \times 10^{23}/mol$
$n$	Natural number
$n_i$	Refractive index of material $i$
$n_{mol}$	Number of mols in a gas volume
$n_{op}$	Refractive index
$n_0$	Refractive index outside of a waveguide
$P$	Power
$P_{el}$	Electric power
$P_{fl}$	Fluidic power
$P_{opt}$	Optimum power
$P_P$	Fluidic power output of a pump
$P_{P,m}$	Maximum fluidic power of a pump
$P_{th}$	Thermal power
$p$	Pressure
$p_{A,U}$	Pressure generated by an actuator supplied with voltage
$p_{A,0}$	Pressure generated by an actuator without voltage supply
$p_a$	Pressure at the outlet



$p_e$	Pressure at the inlet
$p_R$	Reference pressure
$p_S$	Pressure in a gap
$p_{S,m}$	Maximum pressure in a gap
$p_V$	Vapor pressure
$p_0$	Pressure in the surrounding
$p_3$	Pyroelectric coefficient of a piezo in the direction of polarization
$Q$	Quality of an oscillation
$Q_{el}$	Electrical charge
$Q_i$	Charge of a piezo in i-direction
$Q_K$	Heat dissipated by convection
$Q_{th}$	Heat energy
$q$	Quality factor of corrugations
$q_P$	Charge of a particle
$Re$	Reynold's number
$R_B$	Radius of a beam with a circular cross-section
$R_c$	Radius of curvature
$R_{el}$	Electrical resistance
$R_f$	Radius of curvature of a thin film
$R_{fl}$	Fluidic resistance
$R_{fl,Z}$	Fluidic resistance of feed channels
$R_G$	Gas constant = 8.314 J/(mol K)
$R_{Ka}$	Radius of a capillary
$R_M$	Radius of a membrane
$R_{Su}$	Radius of curvature of a substrate
$R_T$	Radius of a particle
$R_{th}$	Thermal resistance
$R_V$	Radius of the boss of a membrane
$R_W$	Radius of a circular wall
$R_Z$	Radius of support and delivery pipes
$R_{Z,max}$	Maximum radius of valve seat at which a pump can generate a flow
$r$	Coordinate in radial direction
$r_K$	Kerr constant
$r_P$	Pockels constant
$S$	Entropy
$s_f$	Scaling factor
$s_{max}$	Maximum deflection of a heated beam clamped at both ends
$s_0$	Deflection
$T$	Temperature
$T_a$	Initial temperature
$T_K$	Critical temperature at which the surface tension of a liquid equals to zero theoretically
$T_0$	Room temperature
$t$	Time

$t_f$	Duration of a delivering pulse of a pump
$t_v$	Switching time of valves
$t_z$	Cycle time
$U$	Voltage
$U_m$	Measurement voltage
$U_{\max}$	Maximum voltage
$U_{\min}$	Minimum voltage
$U_p$	Pull-in-voltage
$U_S$	Circumference length of a valve seat
$U_w$	Wetted circumference length of a fluid
$U_0$	Supply voltage
$u$	Position-dependent part of a wave function
$u_i$	Position-dependent part of a wave function of mode $i$
$u_1$	Static deflection shape of a membrane
$V$	Volume
$V_A$	Volume displaced in one pump cycle
$V_a$	Dead volume of outlet valve
$V_{B,a}$	Volume of a bubble in a pump when the outlet valve is open
$V_{B,e}$	Volume of a bubble when the inlet valve is open
$V_{B,0}$	Volume of a bubble at normal conditions
$V_e$	Dead volume of inlet valve
$V_F$	Volume delivered by one pump cycle
$V_{fl}$	Volume of a liquid
$V_{fl,\min}$	Minimum liquid volume required
$V_L$	Leakage volume per pump cycle
$V_{\text{mol}}$	Volume of a mol = 22.4 L at 101.3 kPa and 22°C
$V_P$	Dead volume of a pump
$V_p$	Potential energy
$V_{p,m}$	Maximum of potential energy
$V_T$	Dead volume
$V_v$	Dead volume of a valve
$V_W$	Volume in the bulged up part of a membrane
$V_0$	Volume at normal conditions
$v$	Mean velocity of a fluid in a capillary or velocity of a mass
$v_a$	Velocity before a change
$v_e$	Velocity a long time after a change
$v_{\max}$	Maximum velocity
$v_k$	Critical flow velocity
$v_w$	Velocity of wall movement
$W$	Work
$W_A$	Energy output of an actuator
$W_E$	Energy input of an actuator
$W_{A,\text{opt}}$	Maximum energy output
$W_C$	Energy stored in a capacitor

$W_H$	Work necessary for valve switching
$W_{th}$	Thermal energy
$w$	Deflection
$w_E$	Deflection of a piezo or an arrangement of piezos generated by the voltage supplied
$w_{E,B}$	Deflection of a bimorph at which no outer force is applied
$w_F$	Deflection of a piezo or an arrangement of piezos generated by outer forces
$w_{F,B}$	Deflection of a bimorph generated by outer forces
$w_i$	Deflection in i-direction
$w_l$	Longitudinal deflection of a beam
$w_{n,m}$	Deflection of the mode n,m
$w_R$	Deflection at the rim
$w_{th}$	Deflection of a thermal actuator when no outer forces are acting
$w_U$	Deflection at which a beam or membrane is snapping over to the opposite side
$w_V$	Pre-deflection of an actuator
$w_x$	Deflection in x-direction
$w_z$	Deflection in z-direction
$w_0$	Deflection at the end of a beam, the center of a membrane, or amplitude of an oscillation
$w_{0,B}$	Deflection of a bimorph
$w_{0,max}$	Maximum deflection
$w_{0y}$	Deflection of a membrane when it bursts under a pressure load
$w_{0,opt}$	Deflection of a piezo at maximum energy output
$x,y,z$	Euclidian coordinates
$\alpha$	Angle
$\alpha_D$	Damping constant
$\alpha_m$	Angle of light propagation of the mode m
$\alpha_{op}$	Optical damping constant
$\alpha_{th}$	Thermal expansion coefficient
$\alpha_S$	Damping by sound emission
$\alpha_T$	Temperature coefficient of an electrical resistance
$\alpha_{Tot}$	Angle of total reflection
$\alpha_t$	Angle of transmitted light
$\alpha_r$	Angle of reflected light
$\alpha_V$	Damping by the viscosity of the surrounding fluid
$\beta$	Angle
$\beta_m$	Additive mass
$\zeta$	$\zeta$ -potential
$\Delta$	Difference
$\Delta b_B$	Change of the width of a beam
$\Delta d_B$	Change of the thickness of a beam
$\Delta L_L$	Change of the length of a conductor path

$\Delta b_L$	Change of the width of a conductor path
$\Delta p$	Pressure difference
$\Delta p_A$	Pressure difference over an actuator
$\Delta p_C$	Capillary pressure on a dielectric
$\Delta p_{EOF}$	Pressure difference generated by an electroosmotic flow
$\Delta p_K$	Capillary pressure
$\Delta p_M$	Pressure difference generated by the elastic force of a membrane
$\Delta p_{max}$	Maximum pressure difference generated by an actuator
$\Delta p_P$	Pressure generated by a pump
$\Delta p_{P,m}$	Maximum pressure generated by a pump
$\Delta p_p$	Pressure difference applied to a capillary from outside
$\Delta p_R$	Pressure difference needed to overcome the friction of a flow
$\Delta p_S$	Pressure drop over a valve seat
$\Delta p_{sf}$	Squeeze-film pressure difference
$\Delta p_V$	Pressure difference over a valve
$\Delta p_y$	Burst pressure of a membrane
$\Delta p_Z$	Pressure difference over feed channels
$\Delta R_{el}$	Change of electrical resistance
$\Delta R_{el,l}$	Resistance change of longitudinal strain gauges
$\Delta R_{el,R}$	Resistance change of radial strain gauges
$\Delta R_{el,T}$	Resistance change of tangential strain gauges
$\Delta R_{el,t}$	Resistance change of transversal strain gauges
$\Delta T$	Temperature difference
$\Delta T_a$	Temperature difference of the wire of an anemometer and the fluid before a change
$\Delta T_e$	Temperature difference of the wire of an anemometer and the fluid a long time after a change
$\Delta V$	Volume change
$\Delta V_A$	Volume change generated by an actuator
$\Delta V_{A0}$	Displaced volume at the beginning of a pump cycle
$\Delta V_B$	Volume change of a bubble
$\Delta V_{max}$	Volume change of an actuator when no counter pressure is applied
$\Delta V_{opt}$	Volume change at maximum energy output
$\Delta \epsilon_r$	Difference of relative permittivity
$\Delta \varphi$	Phase shift
$\Delta \varphi_{n,m}$	Phase shift of the mode n,m
$\epsilon$	Strain
$\epsilon_B$	Strain at the surface of a beam
$\epsilon_b$	Strain in the direction of the width
$\epsilon_d$	Strain in the direction of the thickness
$\epsilon_i$	Strain in i-direction
$\epsilon_l$	Longitudinal strain
$\epsilon_R$	Radial strain
$\epsilon_r$	Relative permittivity

$\varepsilon_T$	Tangential strain
$\varepsilon_t$	Transversal strain
$\varepsilon_{th}$	Thermal strain
$\varepsilon_0$	Absolute permittivity = $8.9 \times 10^{-12}$ (A s)/(V m)
$\eta$	Dynamical viscosity
$\eta_A$	Efficiency of an actuator
$\eta_C$	Efficiency of the Carnot process
$\Phi$	Electrical potential
$\Phi_a$	Flow through an outlet valve
$\Phi_b$	Backflow through a passive valve
$\Phi_F$	Volume flow
$\Phi_P$	Volume flow generated by a pump
$\Phi_{P,m}$	Maximum volume flow generated by a pump
$\Phi_S$	Volume flow over valve seat
$\Phi_{St}$	Slope angle of the mirror of a grid
$\Phi_{th}$	Heat flow
$\Phi_W$	Electrical potential at the surface of a rigid body
$\Phi_Z$	Volume flow through feed channels
$\varphi$	Angle or phase
$\varphi_i$	Phase of vibration mode i
$\Gamma$	Surface tension of a liquid
$\Gamma_K$	Capillary force per circumference length
$\Gamma_{Kap}$	Force due to the capillary “force”
$\kappa$	Coupling constant between neighboring wave guides
$\Lambda$	Vaporization heat of 1 mol
$\Lambda_e$	Effective heat conduction
$\lambda$	Wave length
$\lambda_B$	Blase wave length
$\lambda_D$	Debye–Hückel length
$\lambda_F$	Heat conductivity of a fluid
$\lambda_i$	Frequency parameter
$\lambda_M$	Characteristic length of a membrane
$\lambda_V$	Steam heat
$\lambda_0$	Wavelength in vacuum
$\lambda_1$	Frequency parameter of fundamental mode
$\mu$	Friction coefficient
$\nu$	Poisson’s ratio of a conductor path
$\nu_B$	Poisson’s ratio of a beam
$\nu_M$	Poisson’s ratio of a membrane
$\nu_P$	Poisson’s ratio of a piezo
$\nu_{Su}$	Poisson’s ratio of a substrate
$\Omega$	Angular velocity
$\omega$	Angular velocity or angular frequency
$\omega_{n,m}$	Angular frequency of the mode n,m

$\omega_r$	Angular frequency at resonance
$\pi$	Ludolph's constant "Pi" = 3.141592...
$\pi_i$	Piezoresistive coefficient for straining in i-direction
$\pi_l$	Piezoresistive coefficient longitudinal to a conductor path
$\pi_m$	Mean value of piezoresistive coefficients
$\pi_t$	Piezoresistive coefficient transversal to a conductor path
$\rho_B$	Density of a beam
$\rho_{el}$	Specific electric resistance
$\rho_{el,0}$	Specific electric resistance at the origin of the temperature scale used
$\rho_F$	Density of a fluid
$\rho_M$	Density of a membrane
$\sigma_a$	Stress of a rectangular membrane in direction of its length
$\sigma_B$	Stress of a beam
$\sigma_b$	Stress of a rectangular membrane in direction of its width
$\sigma_D$	Stress due to straining
$\sigma_f$	Stress of a thin film
$\sigma_i$	Frequency parameter
$\sigma_k$	Critical stress
$\sigma_l$	Longitudinal stress
$\sigma_M$	Stress of a membrane
$\sigma_R$	Radial stress
$\sigma_T$	Tangential stress
$\sigma_t$	Transversal stress
$\sigma_{th}$	Stress generated by a temperature change
$\sigma_y$	Yield stress
$\sigma_0$	Residual stress
$\tau_A$	Reaction time of an anemometer
$\tau_p$	Time of pressure rise
$\tau_{th}$	Thermal relaxation time
$\theta$	Angle
$\theta_i$	Wave function of the torsional vibration of a beam at mode i
$\theta_w$	Wetting angle
$\theta_{max}$	Maximum angle at which light can enter or leave a waveguide
$\chi$	Number between 0 and 1 describing the contribution of the Torricelli equation to a pressure loss



# Contents

<b>Introduction</b> .....	1
<b>Scaling Laws</b> .....	3
<b>Elastic Deformations</b> .....	5
<b>Thin Films</b> .....	9
<b>Conductor Paths</b> .....	21
<b>Membranes</b> .....	29
<b>Strain Gauges on Membranes</b> .....	53
<b>Beams</b> .....	65
<b>Vibrations</b> .....	85
<b>Capillaries</b> .....	107
<b>Capacitive Forces</b> .....	131
<b>Piezoelectric Effect</b> .....	137
<b>Thermal Actuators</b> .....	159
<b>Microoptics</b> .....	177
<b>Diffusion</b> .....	191



<b>Microvalves</b> .....	203
<b>Micropumps</b> .....	229
<b>Microdosing</b> .....	251
<b>Analogies of Physical Domains</b> .....	253
<b>Mechanical Devices for Electronics</b> .....	261
<b>Pressure Sensors</b> .....	275
<b>Flow Sensors</b> .....	289
<b>Inertial Sensors</b> .....	303
<b>References</b> .....	311
<b>Index</b> .....	317

# List of Tables

Table 1	Piezoresistive coefficients on {100} silicon wafers .....	25
Table 2	Interrelationship of pressure difference and deflection of a membrane .....	50
Table 3	Area momentums of inertia of beams .....	67
Table 4	Critical force (buckling load) of beams with different boundary conditions .....	72
Table 5	Torsional constants of beams .....	78
Table 6	Equations for the interrelationship of deflection and force or torque of beams and membranes .....	80
Table 7	Frequency parameters of beams clamped at one end .....	96
Table 8	Frequency parameters of beams clamped at both ends .....	98
Table 9	Mass momentums of inertia .....	101
Table 10	Equations of the resonance frequencies of membranes and beams .....	102
Table 11	Hydraulic diameter and flow resistance of microchannels .....	110
Table 12	Equations of pressures acting in capillaries .....	128
Table 13	Equations of deflections, forces, work outputs, and efficiencies of piezos .....	155
Table 14	Correction factors for the arrangement of the piezos .....	156
Table 15	Equations for deflections, forces, work outputs, and efficiencies of thermomechanical actuators .....	163
Table 16	Typical orders of magnitude of properties of microactuators ...	174
Table 17	Pockels constants .....	188
Table 18	Kerr constants .....	188
Table 19	Equations of the energies and forces of a mass fixed to a spring and an electrical circuit consisting of a capacitor, a coil, and a resistance .....	254
Table 20	Devices of different physical domains which are analog to each other .....	256
Table 21	Analog quantities in different physical domains .....	259



# Introduction

This book was developed from lectures held at RWTH Aachen University, Germany and Tsinghua University at Beijing, P.R. China. It may be used as a basis for similar lectures on designing of microsystems. For this purpose, it is recommended to follow the sequence of this book, because it is arranged such that following chapters are building up on previous ones. Students are also strongly recommended to solve the problems included in this book, because this is important for getting familiar with the units and orders of magnitude to be expected in microtechnique. Besides this, often the relevance of the lessons becomes clear much more, when an example calculation shows the importance of the subjects taught.

This book is written such that the reader does not need to know how microsystems are fabricated and what are the big possibilities and limits of microtechnique. When real microsystems are to be designed, a lot of knowledge of the fabrication possibilities is required of cause. This knowledge is not provided in this book, because other literature is available which covers this topic [3–6]. Students may be taught the contents of this book without any prior knowledge of the fabrication processes. However, it is recommended to teach microfabrication first and then the design of microsystems whenever this is possible.

Besides teaching, the purpose of this book is to provide the equations which are needed to calculate the behavior of basic elements and physical effects which are important in microtechnique. For a quick reference, several tables are included, which allow to find the equation needed for a certain problem. At the beginning of this book, there is an index of all tables.

The equations introduced in this book are not restricted to microtechnique. They are all valid in the macroscopic world also. It may be, however, that an effect such as capillary force is less important in macroscopic applications. On the other hand, it is expected that the coherent description of topics such as strain gauges and the piezo-electric effect will also help readers not working on microsystems.

Nowadays, finite element methods (FEM) are available which allow calculating the behavior of macroscopic as well as microscopic structures and elements with high precision. However, FEM do not provide an overview and an understanding of the interrelationships and how to optimize a component for a certain application. Therefore, analytical calculations are desirable which provide an overall understanding of a given problem first, and, after an advantageous way of solving the

problem is envisaged, FEM can be used to find the precise optimum. An approximate description of the problem by an analytical equation will help optimizing by FEM, because the equation gives a good hint where to search for the optimum and which parameters show the largest effect.

The existence of FEM also allows making more rough approximations in analytical calculations, because the calculations are no longer needed to find the exact results. In this book, very rough approximations are accepted achieving an analytical description of the overall behavior of structures such as the buckling up of membranes and beams when the compressive stress exceeds the critical stress. This shows how analytical calculations and FEM nowadays complement each other.

Microsystem technology combines a lot of technical fields such as mechanics, electronics, fluidics, optics, etc. Therefore, the notations of all these fields need to be mixed up, and avoiding confusion of variables is not an easy task. Every variable is assigned a unique notation throughout this book which is found in “Notations and Symbols” section. As a consequence, variables need to be distinguished by subscripts where this is not necessary normally. For example, the Greek letter  $\alpha$  may be used for an angle, a damping constant, the thermal expansion coefficient, and the temperature coefficient of an electrical resistance. These quantities are assigned the distinctive notations  $\alpha$ ,  $\alpha_D$ ,  $\alpha_{th}$ , and  $\alpha_T$ , respectively.

In microtechnique, typically very small structures and shape changes are next to much larger ones. If both small and large structures would be shown up to scale, the smaller ones could not be recognized in general. Therefore, it is usual both in microtechnique and throughout this book to draw smaller structures larger than they are in comparison to surrounding larger ones. The real dimensions are given in the figure caption or the text related to it.

# Scaling Laws

If a cube is to be separated into smaller ones, in such a way, that each edge is cut into two equal pieces (cf. Fig. 1), three cuts through the cube need to be made. Each cut generates a surface area equal to the surfaces of the original cube parallel to the cut. All together, the overall surface is duplicated while the volume remains to be the same. If each edge is cut into  $n$  pieces, the surface is increased by a factor  $n$ . Thus, the ratio of surface to volume is changing by the inverse of the factor  $s_f$  by which the dimensions of a cube (and in general any object) is scaled up or down.

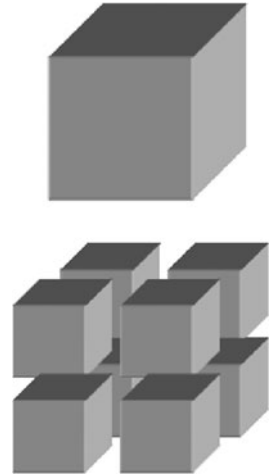
This is the reason why ships are built as large as possible. The building costs are proportional to the hull – the surface of the ship; the driving costs are proportional to the friction with the water – which is proportional to the surface again. But the benefit is proportional to the amount of goods transported – the volume of the ship. Therefore, the costs grow as the surface with the square of the scaling factor  $s_f$  of a ship, while the benefit increases as the volume with the third power of  $s_f$ , i.e., the larger a ship is the more economic it will be. This means also that microships are not a good idea if large amounts of goods are to be transported. However, there are other problems where miniaturization is an advantage.

The simple geometrical fact that the ratio of surface to volume scales with  $s_f^{-1}$  has a lot of consequences in microtechnique in general, and especially for the design of microcomponents and systems. For example, noble metals are widely used in microtechnique because corrosion acts on the large area of microcomponents, while the small volume is used up very quick. If 100  $\mu\text{m}$  of the supporting pillar of a macroscopic bridge is corroded away, nobody will care about that, but if 100  $\mu\text{m}$  of a microstructure are corroded away, it may have vanished. On the other hand, the cost of the small volume of noble metal needed is of no matter.

The increase of chemical reactions is a disadvantage with respect to corrosion, but it is an advantage when a chemical analysis is to be made in a small volume. This is the basis for chemical microreactors and microanalysis chips for DNA recognition and other biological or chemical assays.

The surface to volume ratio is also the reason for a quick exchange of heat for microcomponents. The large surface facilitates heat exchange, while the small volume results in a small heat capacity. The flow of heat, fluids, and electrical current is scaled down when the dimensions are reduced, because the cross-section is diminished with  $s_f^{-2}$  but the length is reduced by  $s_f^{-1}$  only. This shows that

**Fig. 1** Surface increase as a result of the separation of a cube



microtechnique is more than just a reduction of dimensions, but new concepts and new design principles are required.

This becomes especially clear when the scaling down of forces is investigated. All forces that act on the surface of objects become larger in comparison to forces acting on the volume or mass. Therefore, electrostatic force, piezoelectric force, capillary force, and friction are of more importance in microtechnique than gravity, inertia, and magnetism which govern the macroscopic world.

The principal interrelationships that change the nature of things when their overall size is altered are called the scaling laws. They apply not only to effects and forces but also to entire sensors and actuators as described in this book. The scaling laws are the reason why it was necessary to write this book, which describes the distinguished designs required in microtechnique.

# Elastic Deformations

Before the mechanics of basic elements of microtechnique is described, it is necessary to introduce the main parameters which govern the elastic deformation of rigid bodies under the action of forces. There are mainly three parameters: *Stress due to straining, residual stress, and bending.*

Stress due to straining is generated when a force is pulling or pushing at a rigid body. For example, a force  $F$  pulls in longitudinal direction at the free end of a beam which is fixed at its other end (cf. Fig. 2a). This force results in a strain  $\varepsilon_B$  of the beam and a stress  $\sigma_B$ . The stress is the force acting at the beam per cross-section area  $A_B$ . Stress and strain are proportional to each other according to Hooke's law:

$$\frac{F}{A_B} = \sigma_B = E_B \varepsilon_B. \quad (1)$$

The proportionality constant  $E_B$  is a function of the material of the beam and is called Young's modulus.

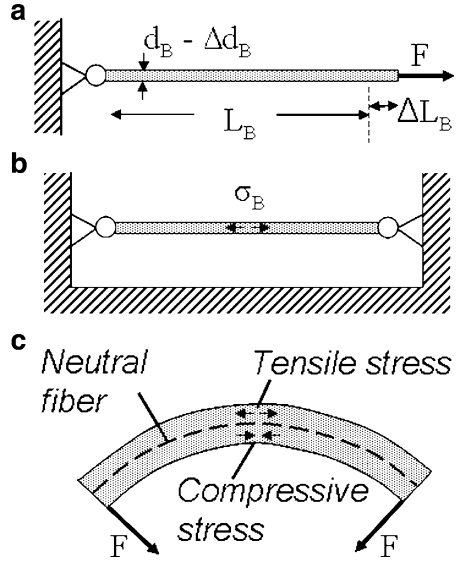
While the length of the beam is increased according to Hooke's law, its diameter is decreased by a factor which is smaller than the longitudinal strain  $\varepsilon_B$  by Poisson's ratio  $\nu_B$ . Thus, the change of the width  $\Delta b_B$  and the thickness  $\Delta d_B$  of the beam are:

$$\Delta b_B = -\nu_B \varepsilon_B b_B \quad \text{and} \quad \Delta d_B = -\nu_B \varepsilon_B d_B. \quad (2)$$

If the force acting at the beam is released again and if the force was not too large, the beam comes back to its original position and there is neither stress nor strain any more. Besides the stress due to straining, a rigid body may show some residual stress also. This occurs only if the body is fixed at more than one part, e.g., a beam fixed at both ends (cf. Fig. 2b). The residual stress in microtechnique often is generated during the fabrication process. For the example of a beam fixed at both ends, it may be assumed that the beam was stretched or compressed to fit between the fixation points. If the residual stress does not exceed certain values and no outer forces are acting, the beam looks the same way as a beam without residual stress, but a significant change is observed when an outer force is acting on the beam. For example, if a force is pulling transversally at the center of the beam, it may be deflected much more with a compressive residual stress and much less with a tensile one.



**Fig. 2** Beams with (a) stress due to straining, (b) residual stress, and (c) bending



Again, if the fixation points are moving or released, the residual stress is changing and the beam shows some movement, e.g., becomes thicker and shorter.

Bending generates tensile stress and strain on one side of a body and compressive ones on the opposite side (cf. Fig. 2c). The gradient of the stress over the thickness of the body results in a bending moment, which tends to bring the body back into its original position when outer forces are taken away again. The integral of the stress over the cross-section is zero when only bending is involved. So, bending does not alter the length of, e.g., a beam. There is an infinitesimally thin layer in every bent body at which the stress is zero. This is called the neutral fiber. The neutral fiber is strained, however, when a force or a force component is acting longitudinal to the beam. Beams clamped at one end and loaded only in transversal direction are elastically deformed only by bending, because the neutral fiber cannot be strained that way.

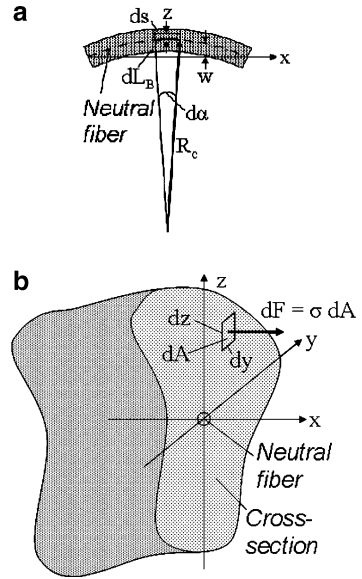
The strain  $\varepsilon$  of a beam in axial direction generated by bending is the change of an infinitesimal length  $dL_B$ , cf. Fig. 3a:

$$\varepsilon = \frac{ds - dL_B}{dL_B} = \frac{(R_c + z)d\alpha - R_c d\alpha}{R_c d\alpha} = \frac{z}{R_c}. \quad (3)$$

In this equation,  $R_c$  denotes the radius of curvature of the beam. The radius of curvature of a function  $w(x)$  can be calculated, in general, by [21]:

$$R_c = \pm \frac{\left(1 + (\partial w / \partial x)^2\right)^{3/2}}{\partial^2 w / \partial x^2}. \quad (4)$$

**Fig. 3** Calculation of strain and bending moment of a beam



For small slopes of the curve (if the x-axis is parallel to the beam axis), the first derivative can be neglected compared with one, and (3) becomes:

$$\varepsilon = -z \frac{\partial^2 w}{\partial x^2}. \tag{5}$$

The bending moment  $M$  acting at the beam is the integral of the product of the force  $dF$  acting on the infinitesimal surface element  $dA$  of the cross-section of the beam and the distance  $z$  to the axis around which the beam is bending (the y-axis in Fig. 3b):

$$M = \int_A z dF = \int_A z \sigma dA. \tag{6}$$

In this equation, the stress according to Hooke's law (1), and (5) is inserted:

$$M = \int_A z \sigma dA = - \int_A z^2 E_B \frac{\partial^2 w}{\partial x^2} dA = -E_B \frac{\partial^2 w}{\partial x^2} \int_A z^2 dA = -E_B I \frac{\partial^2 w}{\partial x^2}. \tag{7}$$

In this equation,  $I$  denotes the area momentum of inertia which is defined by:

$$I = \int_A z^2 dA. \tag{8}$$

The area momentum of inertia of cross-sections important in microtechnique are listed in Table 3 on page 67.

# Thin Films

Thin films are an important basic element of microsystem design. They are used for masking in etching processes, as a diffusion stop layer, and as functional elements such as electrical conductors, membranes, and beams. Thin films may consist of nearly every rigid material. Typical examples are metals, polymers, oxides, and nitrides. The thickness of thin films typically is in the range of 50 nm – 10  $\mu\text{m}$ . The lower limit is due to the problem that layers with an average thickness of less than 50 nm hardly are made homogeneously, because they tend to form separated clusters. The upper limit is a kind of convention. Films that are thicker than 10  $\mu\text{m}$  are no longer considered to be thin in microtechnology, and they cannot be generated easily by processes such as sputtering and evaporation but need to be produced by, e.g., electroplating.

When dealing with thin films, stress is an important parameter which may make the difference between success and a flop of a newly developed microsystem. That is why stress control is important (or even a design which works at any stress level). This chapter describes how thin film stress affects their behavior, what are the consequences for parts fixed to the film, and how the stress can be altered.

It is almost impossible to deposit a thin film onto a substrate without any residual stress. The only exception is epitaxial growth. The reason for this is that the molecules of the deposited layer walk around on the substrate surface until an energetically low position is found. The mismatch in the crystalline structure of substrate and thin film material results in a strain of the crystal lattice of the thin film, and, therefore, generates some stress.

If a polymer layer is deposited by a process such as spin-coating or just painting, a different mechanism is working. A solvent is evaporating from the thin film and its dimension is reduced. As a consequence, some tensile stress is generated in the thin film.

Residual stress bends the substrate a little bit. Typically, the deflection by bending of a silicon wafer with a diameter of 100 mm and a thickness of 500  $\mu\text{m}$  is on the order of 100  $\mu\text{m}$  when the thickness of the thin film is 100 nm. On the macroscopic scale, this is of no importance. When a bridge is painted, nobody cares about the change in shape entailed with this. On the microscopic scale, however, this is an effect which needs to be considered.

Substrate bowing is used to measure residual stress  $\sigma_0$  of thin films. When the thickness  $d_f$  of the film is much less than the thickness  $d_{Su}$  of the substrate, the residual stress can be calculated approximately with Stoney's equation [9]:

$$\sigma_0 = \frac{E_{Su}}{6(1 - \nu_{Su})R_{Su}} \frac{d_{Su}^2}{d_f}. \quad (9)$$

As can be seen from (9), it is not necessary to know the elastic properties of the thin film which are hard to measure in general. It is sufficient to know Young's modulus  $E_{Su}$ , Poisson's ratio  $\nu_{Su}$ , and the radius of curvature  $R_{Su}$  of the substrate. The radius of curvature of the substrate is easily measured with a surface profiler.

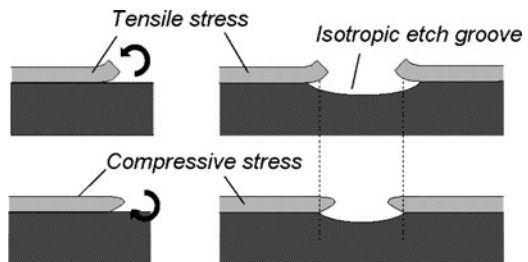
Thin films with a tensile residual stress tend to peel off the substrate [10], while compressive stress facilitates adhesion. As indicated in Fig. 4 (The figures are not up to scale allowing to recognize small dimensions next to larger ones.), a tensile stress pulls the upper part of the thin film away from the rim generating a bending moment which peels the film off the substrate. Therefore, in the light of good adhesion, compressive stress is desirable in thin films.

The stress has also an influence on the etching under a thin film used as a mask. As shown in Fig. 4, the curling down of a mask with compressive stress reduces the undesired etching under the mask while a tensile stress results in an enhanced sideward etching.

A thin film may also curl due to a *stress gradient*. Such gradients occur often when a thin film is deposited on a substrate. As described above, the mismatch in the crystalline structure of substrate and thin film material results in a strain in the crystals of the thin film generated. When the crystal growth is continued, the mismatch disappears and the natural lattice of the thin film is formed. This results in a residual stress with an absolute value getting smaller with increasing distance from the substrate. Figure 5 shows what happens if a sacrificial layer beneath a thin film is etched away when the film initially was under compressive stress decreasing with the distance from the substrate. The beams in the left part of Fig. 6 are curling up because of a gradient in their residual stress.

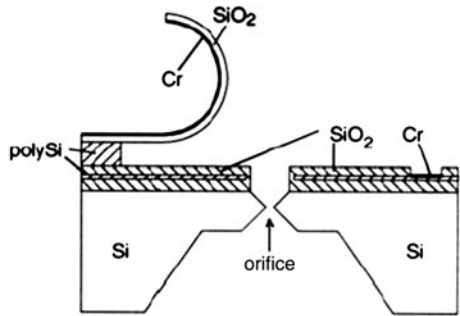
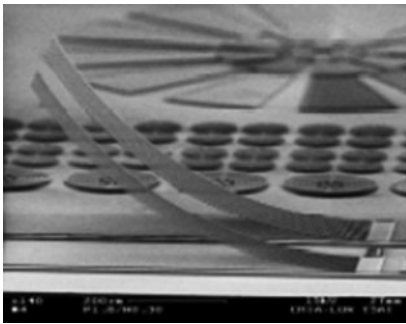
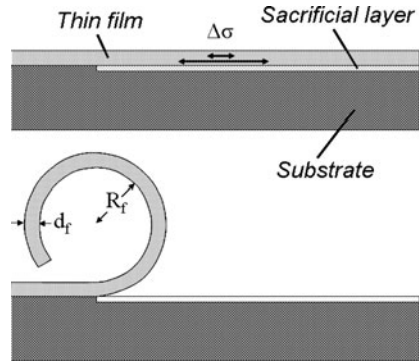
The radius  $R_f$  of curvature of a curling thin film with thickness  $d_f$  and Young's modulus  $E_f$  and a difference in stress  $\Delta\sigma$  can be calculated with the following equation:

$$R_f = \frac{d_f E_f}{\Delta\sigma}. \quad (10)$$



**Fig. 4** Thin film on a substrate with tensile and compressive stress, respectively

**Fig. 5** Thin film on a substrate with a stress gradient after deposition (*top*) and after etching the sacrificial layer (*bottom*)



**Fig. 6** Thin films bowing up (*left*) [7] and two thin films curling up and down as a design element in a microvalve (*right*) [8]. © [1997] IEEE

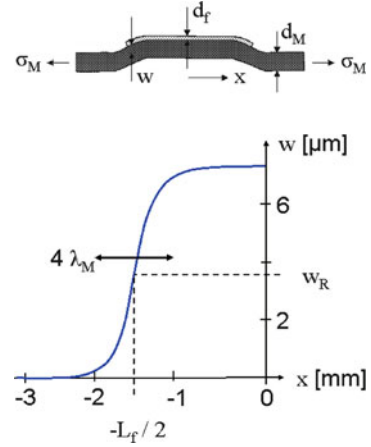
Thin films tend to deform membranes as well. If a membrane with tensile stress carries a thin film, it is very unlikely that the thin film shows no stress. Therefore, it has to be expected that bending moments in the film will deform the membrane. The upper part of Fig. 7 shows a thin film with compressive stress on a membrane which has tensile stress, because it is fixed at a frame (frame not shown in Fig. 7). The bending moments produce a deflection of the membrane in the vicinity of the rim of the thin film. Far away from the rim, there is nearly no bending of the membrane neither in the part uncovered with the thin film nor in the covered part.

If the thickness  $d_f$  of a thin film is much less than the thickness  $d_M$  of the membrane, the following equations can be used to calculate the deflection  $w$  of the membrane with a tensile stress  $\sigma_M$  as a function of the distance  $x$  from the center of the thin film, its length  $L_f$ , and its stress  $\sigma_f$  [11].

The deflection of the part of the membrane covered with the thin film ( $|x| < \frac{1}{2} L_f$ ) is described by:

$$w = \frac{1}{2} \frac{\sigma_f d_f}{\sigma_M d_M} (d_f + d_M) \left( 1 + \left[ \frac{1}{1 + \coth(L_f / (2 \lambda_M))} - 1 \right] \frac{\cosh(x / \lambda_M)}{\cosh(L_f / (2 \lambda_M))} \right) \quad (11)$$

**Fig. 7** Thin film on a substrate with compressive stress. *Top*: Schematic drawing; *bottom*: Deflection calculated with (11–14)



and the deflection of the uncovered part ( $|x| > \frac{1}{2} L_f$ ) is given by:

$$w = w_R e^{(-|x| + L_f/2)/\lambda_M} \quad (12)$$

with:

$$w_R = \frac{1}{2} \frac{\sigma_f d_f}{\sigma_M d_M} (d_f + d_M) \frac{1}{1 + \coth(L_f/(2 \lambda_M))}. \quad (13)$$

The parameter  $\lambda_M$  in (11–13) is defined as:

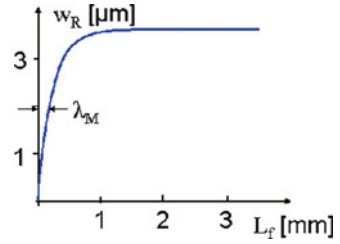
$$\lambda_M = \frac{d_M}{2 \sqrt{3}} \sqrt{\frac{E_M}{\sigma_M(1 - \nu_M^2)}}. \quad (14)$$

The lower part of Fig. 7 shows the deflection of a membrane with a thickness of 5  $\mu\text{m}$ , Young's modulus of 120 GPa, Poisson's ratio of 0.3, and a tensile stress of 10 MPa as calculated with (11–14). It is bent by a 1- $\mu\text{m}$  thick thin film with a length of 3 mm. The parameters  $\lambda_M$  and  $w_R$  are 234 and 3.6  $\mu\text{m}$ , respectively. As shown in Fig. 7, the bending of the membrane is restricted to a stripe with a width of approximately  $4\lambda_M$  and the total deflection is  $2w_R$ . Note that the total deflection is larger than the thickness of the membrane!

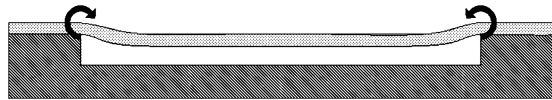
Figure 8 displays the deflection parameter  $w_R$  calculated with (13) as a function of the length of the thin film. It is clearly visible that the total deflection due to the stress in a thin film on a membrane may be reduced substantially when the length of a continuous film is designed smaller than  $\lambda_M$ . So, an undesired deflection may be avoided by separating the thin film into parts smaller than  $\lambda_M$ .

A membrane will be bent also, if it is fixed with tensile stress to a frame on one side only and if there is no thin film on the membrane (cf. Fig. 9). The upper part of the membrane is relaxing a little bit by moving from the frame to the free-span membrane. Thus, a bending moment is generated which pushes the membrane down to the side of the frame.

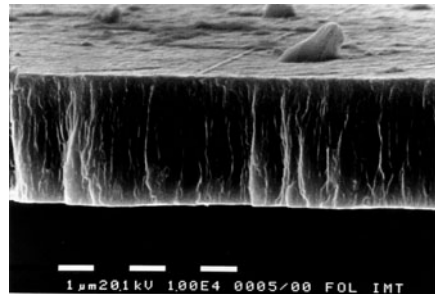
**Fig. 8** Deflection parameter  $w_R$  calculated with (13)



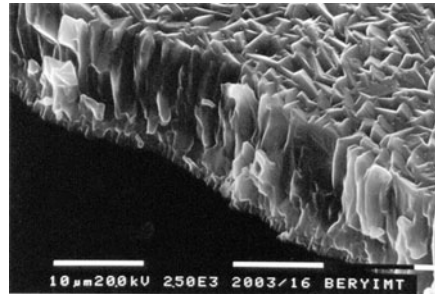
**Fig. 9** Deflection of a membrane stretched with tensile stress over a frame



**Fig. 10** SEM of a titanium membrane, 2.7  $\mu\text{m}$  in thickness. (Courtesy of Karlsruhe Institute of Technology, KIT)



**Fig. 11** SEM of a beryllium membrane, 10  $\mu\text{m}$  in thickness. (Courtesy of Karlsruhe Institute of Technology, KIT)



Figures 10 and 11 show breaking edges of membranes sputtered from titanium and beryllium, respectively, as observed with a scanning electron microscope (SEM). It is clearly seen that the *crystalline structure* of these membranes is different. As a consequence, the properties of these membranes are different. Crystal growth during sputtering the beryllium started from the bottom in Fig. 11, and it is seen that at the bottom there are the smallest crystallites. The speed of growth is a function of the orientation of the crystal lattice. Some of the crystallites in the bottom layer are orientated such that they show a quicker growth and they grow over their neighbors. Therefore, the properties of such a thin film become more anisotropic when it is thicker.

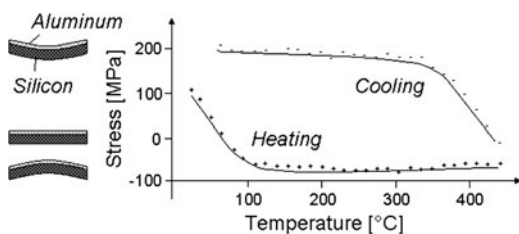
In Fig. 10, the crystallites are very small and homogenous. The reason for this is that the crystalline structure was changed by exposing the thin film to high temperatures after sputtering. This has caused *recrystallization*, i.e., the atoms of the material started to move at elevated temperatures and to change their position from one crystal to the other. This process is lent by the reduction of stress. Thus, the recrystallization continues until a certain stress level is achieved which does not induce further recrystallization.

This became evident in an experiment performed by Flinn [12]. He used silicon wafers on which a thin metal layer had been sputtered. The films showed tensile stress, and, therefore, the wafers were bent. Flinn used the curved thin film as a hollow mirror. The focal length of this mirror was a measure of wafer bending and thin film stress. This way, the stress of the thin film could be measured through the window of an oven and the change of the stress was recorded as a function of temperature. Figure 12 shows the result of this experiment for aluminum films on silicon wafers.

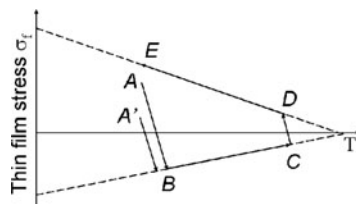
The experiment started at a thin film stress of approximately 100 MPa. Then the temperature was raised and the stress was reduced according to the larger thermal expansion of aluminum compared with silicon. As temperature continued to increase the thin film got compressive stress. At approximately 100°C and a stress of -70 MPa, recrystallization started and the compressive stress no longer changed with raising temperature or even reduced. At 450°C, heating was turned off and cooling down of the sample started. Again the stress was changing according to the difference in thermal expansion of aluminum and silicon until a certain tensile stress was reached and recrystallization reduced further stress enhancement.

A theoretical model of this experiment is shown in Fig. 13: During heating the stress follows the line from A to B. At B, a theoretical line is reached which marks the stress level as a function of temperature at which recrystallization starts. From B to C the stress follows this line. When cooling starts there is no longer enough stress for recrystallization and the stress follows the difference in thermal expansion until the tensile stress at D gets so large that the stress follows another line which marks the start of recrystallization.

**Fig. 12** Thin film stress as a function of heating and cooling (reproduced from [12])



**Fig. 13** Theoretical model of the experiment in Fig. 12 [12]





This experiment shows how the stress in thin films can be altered by a temperature cycle. When thin films are sputtered or evaporated onto substrates the conditions differ a bit as a function of the position in the machine. Therefore, thin film stress is not everywhere the same, neither on neighboring substrates in a machine nor on a single substrate. In Fig. 13, it can be seen how the stress of a whole batch may be adjusted to a common level. A thin film starting at  $A'$  instead of  $A$  will end up at the same stress at  $E$  as all the other films with different initial stress.

A stress gradient is reduced or vanishes after temperature cycling. Curling up of the thin films shown in the left part of Fig. 6 (on page 11) might have been avoided with a proper temperature cycle before release from the substrate. The titanium membrane shown in Fig. 10 before separation from the substrate underwent a temperature cycle for 30 min at 450°C. As a result, the membrane stress was changed from compressive to tensile and no more stress gradient was found. The figure shows also the homogeneous crystalline structure generated this way.

The crystalline structure affects other properties of thin films also. One example is the electrical resistance. Usually, conductor metal paths are exposed to a temperature cycle to avoid a later change of the resistance when it may get heated during use.

The strength of thin films is also a function of their crystalline structure. Thin films which have been sputtered or evaporated typically show a larger strength than the casted material. Recrystallization will reduce this strength.

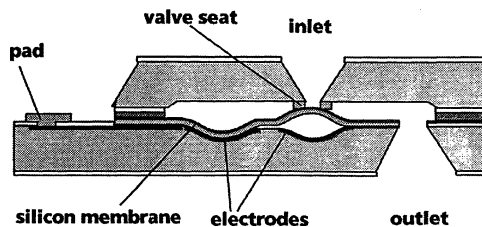
The density of thin films after deposition is approximately 10% less than the density of the same material after casting. Recrystallization increases the density.

Other properties of thin films, which are important for their application, are adhesion to the substrate and other layers, diffusion of molecules and atoms into and through the film, wetting by liquids, chemical reactivity, and thermal conductivity.

## Exercises

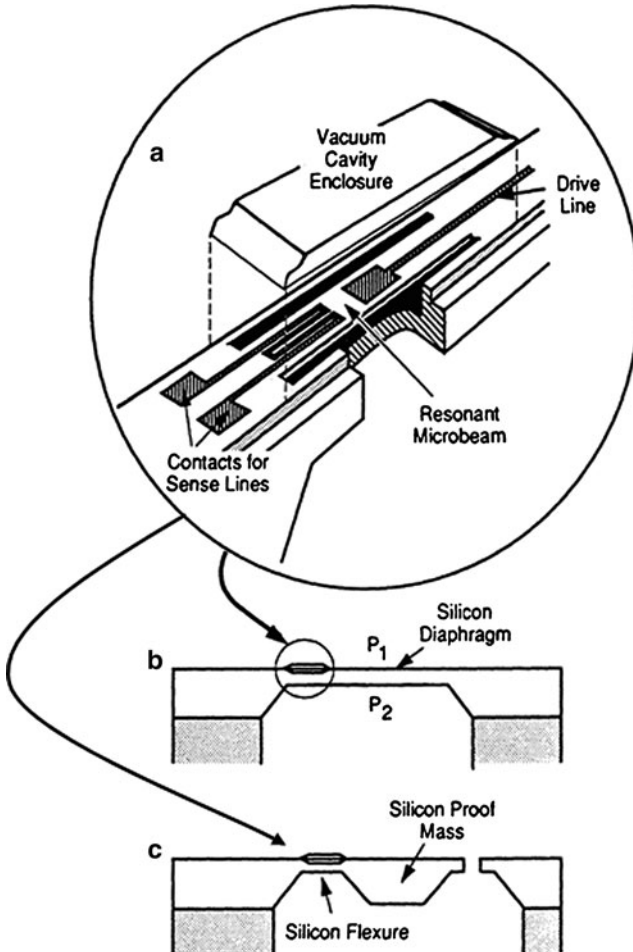
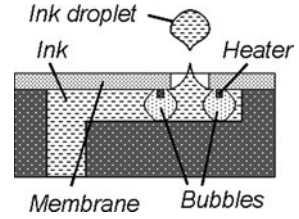
### *Problem 1*

In Figures E1–E4, you find four typical applications from microtechnique. In these applications, basic elements have been employed which you have got to know in the lecture. Find out which are these basic elements.



**Fig. E1** Schematic setup of a bistable microvalve [13]

**Fig. E2** Cross-section of the nozzle of an ink-jet printer



**Fig. E3** Schematic setup of a pressure sensor (b) and an acceleration sensor (c). (Reprinted from [14] with permission from Elsevier)

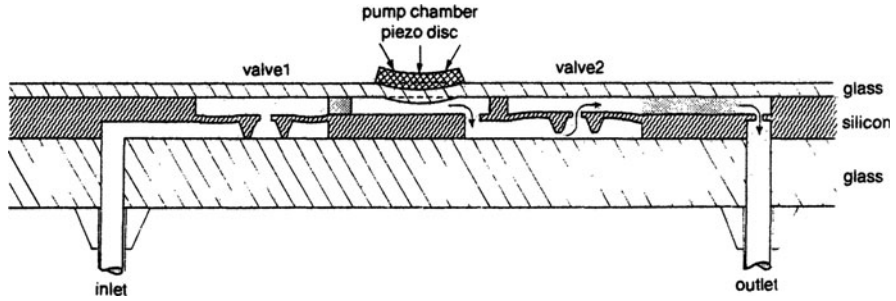


Fig. E4 Silicon micropump [15]

**Problem 2**

Why is it important to know the mechanical stress of a layer? Which meaning has it for the development of a microtechnical system?

**Problem 3**

You got to know Stoney’s equation in the lecture, which you can use to calculate the stress of a thin film as a function of the radius of curvature. Consider which experimental methods you could use to measure the curvature of the substrate. There are several possibilities.

How could you measure the thickness of a thin layer?

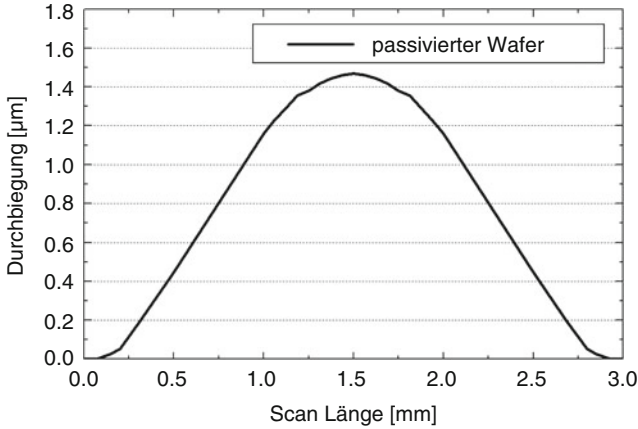
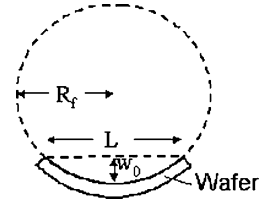
**Problem 4**

Let us assume that you are able to measure the vertical deflection of the substrate with the method suggested by you. The geometrical situation is approximated by a circle (cf. Fig. E5). If the length of a chord of the circle is  $L$  and its maximum distance to the circle is  $w_0$ , the radius can be calculated or approximated, respectively, with the following equation:

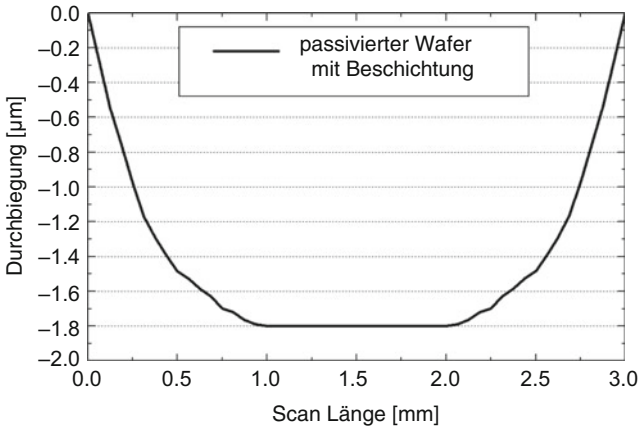
$$R_f = \frac{w_0}{2} + \frac{L^2}{8 w_0} \approx \frac{L^2}{8 w_0}.$$

Figure E6 displays the curvature of a passivated silicon wafer (diameter 100 mm and thickness 0.55 mm) without any metallic layer as measured by Thomas et al.

**Fig. E5** Chord of the circle of curvature



**Fig. E6** Curvature of the uncovered passivated silicon wafer [16]. © [1988] IEEE



**Fig. E7** Curvature of the covered passivated silicon wafer [16]. © [1988] IEEE

[16]. After this measurement, the wafer was sputtered first with 5 nm titanium and then with 500 nm tungsten. The curvature of the covered wafer is shown in Fig. E7.

- (a) Obviously, the passivated but uncovered wafer shows some stress already. What is the reason for this curvature?
- (b) Please, calculate the (mean) stress in the coating of the silicon wafer.

Hint: Do not get irritated because the curvature is no exact circle. This cannot be expected from a real measurement.

Young's modulus of silicon	133 GPa	Poisson's ratio of silicon	0.28
----------------------------	---------	----------------------------	------

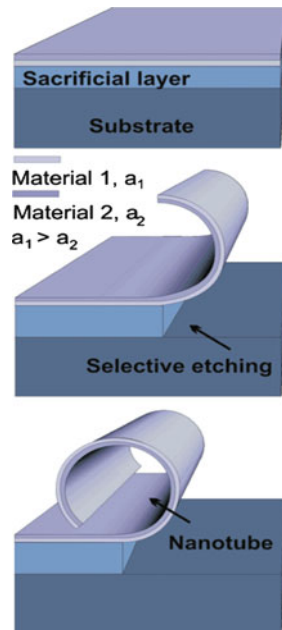
**Problem 5**

In surface micromachining, a sacrificial layer in the near of the surface is removed to partly separate movable parts from the substrate. A wafer is first coated with a sacrificial layer from SiO<sub>2</sub> and then with a thin titanium layer. The sacrificial SiO<sub>2</sub> layer is etched away with hydrofluoric acid.

- (a) A thin layer may curl after removing its linkage to the substrate. This effect is due to a stress gradient in the layer. Calculate the radius of curvature of titanium films with different thicknesses and a stress difference of 300 MPa. The thicknesses of the titanium films are 10 μm, 5 μm, and 200 nm.

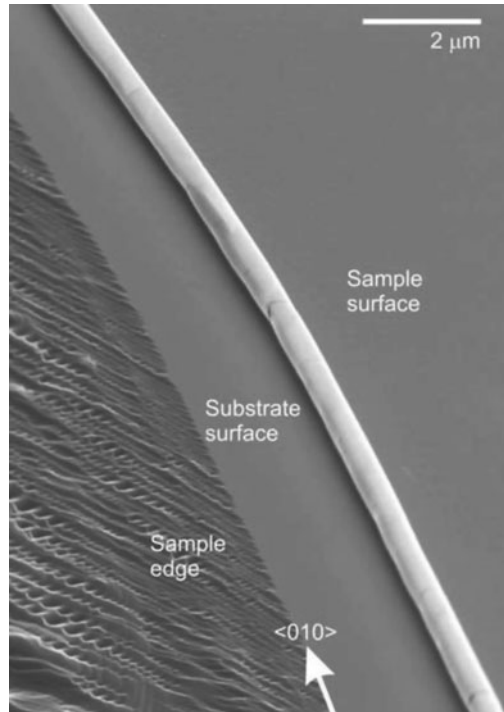
Young's modulus of titanium	120 GPa
-----------------------------	---------

- (b) You may learn from your calculations that the radii of curled films can become very small. With a modified approach at Max-Planck Institute for rigid body research in Stuttgart so-called nanotubes were generated (cf. Figs. E8 and E9). The stress gradient was generated in this case by deposition of two layers with



**Fig. E8** Schematic drawing of the curling of a nanotube consisting of two thin layers. The upper layer shows a smaller lattice distance than the lower one. Courtesy of [17]

**Fig. E9** SEM of a nanotube with a diameter of 530 nm. Courtesy of [17]



a thickness of approximately 60 nm with different stresses. How large was round about the difference in stress of the two layers before the sacrificial layer had been dissolved? The diameter of the nanotube is 530 nm. Assume that the mean Young's modulus of the two layers is 106.5 GPa. Assume that the effective distance of the two layers equals the distance of their centers.

# Conductor Paths

Conductor paths are either thin films on a substrate patterned to get the shape of the desired paths or the substrate itself is made conductive along the desired paths. The latter is obtained by doping of a semiconductor. Doping is obtained by techniques such as diffusion of chemical species such as boron or phosphorus out of a gas phase into a semiconductor along the paths where a conductor is to be formed or ions are implanted there. For more details on doping of semiconductors, refer books on micromachining such as refs. [3–6].

Conductor paths are used to establish electrical connections between components and are employed as strain gauges measuring strain. Typical thicknesses of conductor paths range from 50 nm to 5  $\mu\text{m}$ . The lower limit is defined by the ability to form a continuous layer as described in the beginning of the previous chapter and the upper limit is set by the effort necessary to generate thicker layers with the processes used in microtechnique. The lateral dimensions of conductor paths in electronics fabrication nowadays achieve 40 nm or even less. Wiring in microsystems does not require such small dimensions. Typically, the width of conductor paths is in the range from a few micrometers up to some millimeters.

This chapter describes how the resistance of a conductor path is affected by strain and temperature providing the basics for designing *strain gauges*.

The resistance  $R_{\text{el}}$  of a conductor path with length  $L_L$ , thickness  $d_L$ , width  $b_L$ , and specific resistance  $\rho_{\text{el}}$  can be calculated with the following equation:

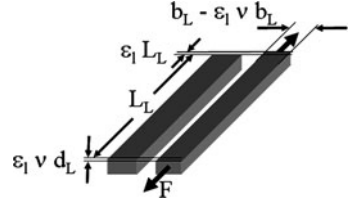
$$R_{\text{el}} = \rho_{\text{el}} \frac{L_L}{b_L d_L}. \quad (15)$$

If this conductor path is strained in the direction of its length by  $\Delta L_L = L_L \varepsilon_l$ , width by  $\Delta b_L = b_L \varepsilon_b$ , and thickness  $\Delta d_L = d_L \varepsilon_d$ , the resistance changes according to the change in geometry:

$$\Delta R_{\text{el}} = \frac{\partial R_{\text{el}}}{\partial L_L} \Delta L_L + \frac{\partial R_{\text{el}}}{\partial b_L} \Delta b_L + \frac{\partial R_{\text{el}}}{\partial d_L} \Delta d_L = R_{\text{el}}(\varepsilon_l - \varepsilon_b - \varepsilon_d). \quad (16)$$

If a unidirectional strain  $\varepsilon_l$  is applied to this conductor path in the direction of its length (cf. Fig. 14), it becomes longer by the factor  $(1 + \varepsilon_l)$ , and if there are no

**Fig. 14** Geometrical changes of a conductor path due to a longitudinal force



other boundary conditions, i.e., the conductor path is not fixed to a substrate, it will shrink in the two directions perpendicular to the straining by a factor  $(1 - \nu\epsilon_1)$  with  $\nu$  denoting Poisson's ratio. So, the width and the thickness of the conductor path become smaller when it is strained in the direction of its length. The resulting resistance change is shown in (17):

$$\Delta R_{el} = \frac{\partial R_{el}}{\partial L_L} \Delta L_L + \frac{\partial R_{el}}{\partial b_L} \Delta b_L + \frac{\partial R_{el}}{\partial d_L} \Delta d_L = R_{el}(\epsilon_1 + \nu\epsilon_L + \nu\epsilon_L) = R_{el}\epsilon_1(1 + 2\nu). \quad (17)$$

Thus, the resistance of a conductor path gets larger by straining when it becomes longer and when it becomes narrower and/or thinner.

This is the resistance change due to the geometrical effect. Besides this, the specific resistance of the conductor material is a function of strain and temperature  $T$ . Equation (16) becomes now:

$$\Delta R_{el} = R_{el}(\epsilon_1 - \epsilon_b - \epsilon_d) + \frac{\partial R_{el}}{\partial \rho_{el}} \left( \frac{\partial \rho_{el}}{\partial \epsilon_1} \epsilon_1 + \frac{\partial \rho_{el}}{\partial \epsilon_b} \epsilon_b + \frac{\partial \rho_{el}}{\partial \epsilon_d} \epsilon_d + \frac{\partial \rho_{el}}{\partial T} \Delta T \right). \quad (18)$$

The specific resistance is nearly a linear function of temperature which is described by the temperature coefficient  $\alpha_T$ :

$$\rho_{el} = \rho_{el,0}(1 + \alpha_T T), \quad (19)$$

where  $\rho_{el,0}$  denoting the specific resistance at the origin of the temperature scale used. Every conductor material shows some resistance change as a function of temperature. Therefore, this is an effect which needs to be taken into account.

The derivative of the specific resistance with respect to the strain  $\epsilon_i$  in direction  $i$  is described by the *gauge factor*  $K$ :

$$\frac{\partial \rho_{el}}{\partial \epsilon_i} = \rho_{el} K_i. \quad (20)$$

Inserting (20) and (19) in (18) yields:

$$\Delta R_{el} = R_{el}(\epsilon_1 - \epsilon_b - \epsilon_d) + R_{el}(K_1 \epsilon_1 + K_b \epsilon_b + K_d \epsilon_d + \alpha_T \Delta T). \quad (21)$$



The gauge factor of metals is comparatively small and usually neglected in calculations. Thus, for metal conductors, only the geometrical and temperature effect is taken into account. The gauge factor of monocrystalline silicon is on the order of 90 and for polycrystalline silicon in the range of 20–40 while the geometrical effect is approximately 1–2. Therefore, if the resistance change of silicon is to be calculated, the geometrical effect is neglected in general.

Sometimes the change of the specific resistance is expressed as a function of the stress  $\sigma_f$  of the conductor path. Then, the terms  $(K_i \epsilon_i)$  in (21) are exchanged against  $(\pi_i \sigma_i)$ , where  $\pi_i$  denotes the *piezoresistive coefficient* in direction  $i$ . The gauge factors and piezoresistive coefficients are measured in experiments not corrected for the transverse strain which occurs when the sample is strained longitudinal. Thus, these material properties already include the effect of transverse strain and there is no stress  $\sigma_d$  in the direction of the thickness:

$$\Delta R_{el} = R_{el}(\pi_l \sigma_l + \pi_b \sigma_b + \alpha_T \Delta T). \quad (22)$$

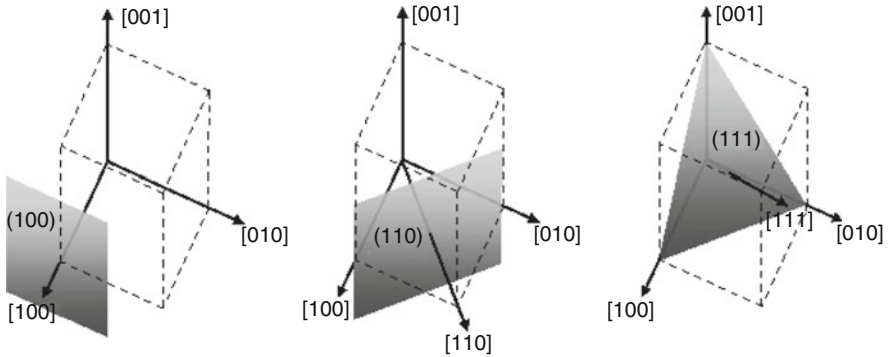
Pure silicon is a semiconductor, i.e., it is insulating at room temperature. It is necessary to add other atoms to obtain electrical conductivity. This is achieved by a process called doping, during which other atoms such as boron or phosphor are brought into the silicon. If the kind of atoms added have less than four electrons in their outer electron shell such as boron, this process is called p-doping, and if there are more than four electrons in the outer shell it is n-doping.

In *monocrystalline materials* such as silicon wafers, the gauge factor and the piezoresistive coefficient are a function of the orientation relative to the crystal lattice. The orientation relative to the lattice is described by the so-called *Miller indices*. Silicon has a cubic lattice, and planes in the lattice are described by their intersection points with the coordinate axes parallel to the lattice. The unit along the coordinate axes is the grid constant which is the distance between two lattice points. The Miller indices of a certain plain through the lattice are the reverse of the distances from the origin of the intersection points with all three axes. So, the Miller indices (111) indicate a plane which shows intersection points with all coordinate axes at the distance of one grid constant (cf. Fig. 15).

Intersections at negative coordinate values are denoted by a minus sign above the corresponding index, e.g. ( $\bar{1}\bar{1}1$ ). However, a minus sign simply mirrors the plane, and, therefore, results in a physically identical plane. In a cubic lattice also, a rotation by 90° results in a physical identical plane. Therefore, the planes (100), (010), and (001) are all the same. It is usual to use a curly brace {100} instead of a parenthesis to denote one of these physically identical planes.

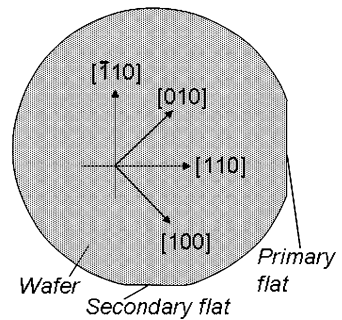
Directions in the lattice are denoted by the Miller indices of the plane perpendicular to the direction. A direction is distinguished from a plane by showing the Miller indices in square brackets such as [100] or [110], instead of parenthesis. In this book, physical identical directions such as [100] and [010] are denoted by sharp brackets  $\langle 100 \rangle$ .

For practical applications in microtechnique, typically the three planes {100}, {110}, and {111} are important only. These planes are shown in Fig. 15.



**Fig. 15** Directions and plains in the cubic silicon lattice

**Fig. 16** Directions on the surface of a (100)-silicon wafer

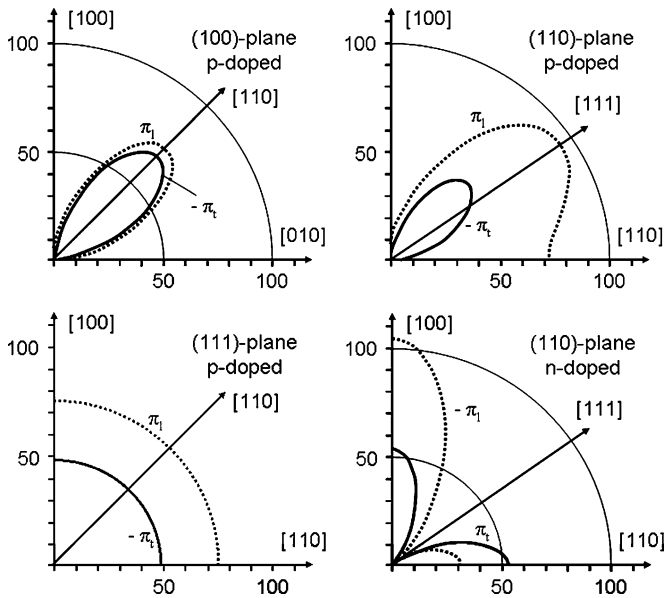


In a silicon factory, the monocrystals are sawn into plates, the so-called silicon wafers. The surface of a wafer is parallel to a certain plain of the crystal lattice such as (100). The directions on the surface of the wafer are indicated by a so-called flat. This is a secant at the circular wafer as shown in Fig. 16. The flat points in a certain direction, e.g. [110].

As described above, the change of the specific resistance of monocrystalline silicon is a function of the orientation on the wafer. Thus, the piezoresistive coefficient of a conductor path is defined by the direction, e.g. [110] on a (100)-wafer, in which it is orientated and whether it is strained parallel or perpendicular. The piezoresistive coefficients were drawn by Kanda in certain graphs [18]. Four of these graphs are shown in Fig. 17.

The piezoresistive constants of some important crystal orientations are shown in Table 1.

The piezoresistive constants are a function of doping also. As shown in Fig. 17, (110)-wafers show completely different piezoresistive constants as a function of orientation, when they are n-doped instead of p-doped. Besides this, the effect is a function of the number of doping atoms per  $\text{cm}^3$  as shown in Fig. 18. Lower-doped

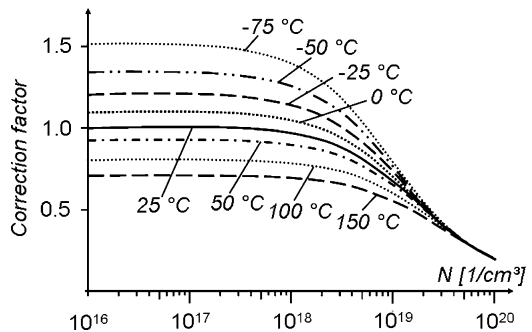


**Fig. 17** Piezoresistive coefficient of (100), (110), and (111) wafers as a function of doping and the orientation in the silicon lattice. Note the different sign of the constants for unidirectional stressing longitudinal  $\pi_l$  and transversal  $\pi_t$  to the conductor path. The constants are shown as the distance from the origin in the unit  $10^{-11}/\text{Pa}$ . The graphs are reproduced from [18]. © [1982] IEEE

**Table 1** Piezoresistive coefficients on {100} silicon wafers doped less than  $10^{18}/\text{cm}^3$

Doping	Orientation	$\pi_l$ ( $10^{-11}/\text{Pa}$ )	$\pi_t$ ( $10^{-11}/\text{Pa}$ )
p	<100>	6.6	-1.1
p	<110>	72	-66
n	<100>	-102.2	53.4
n	<110>	-31	-18

**Fig. 18** Correction factor of piezoresistive constants for doping concentration and temperature. The graph is reproduced from [18].



© [1982] IEEE

silicon is more sensitive to straining than higher-doped. However, lower doping is entailed with a larger temperature effect. Therefore, a doping of  $10^{20}$  atoms per  $\text{cm}^3$  and higher appears to be more favorable for applications where a cross-sensitivity to temperature changes is not desirable.

## Exercises

### Problem 6

Your colleague is proud on his new developed microbalance. It employs the piezoresistive effect to measure small masses. The sensitive element of this balance is a rectangular upright polysilicon beam, 2 mm in length with a cross-section of  $0.25 \times 0.25 \text{ mm}^2$ . At the upper end of the beam, there is a small plate mounted carrying the small samples. The weight of the samples is determined from the change of the electrical resistance along the beam.

- Your colleague measures the weight of a fly in his laboratory. He obtains a resistance change of  $20 \text{ m}\Omega$ . What is the mass of the fly?
- After lunch at the sunny terrace, your colleague shows his new microbalance to his fellow colleagues. The same fly as before is used as a sample. At the terrace the temperature is  $35^\circ\text{C}$ , while in the laboratory there were  $25^\circ\text{C}$ .

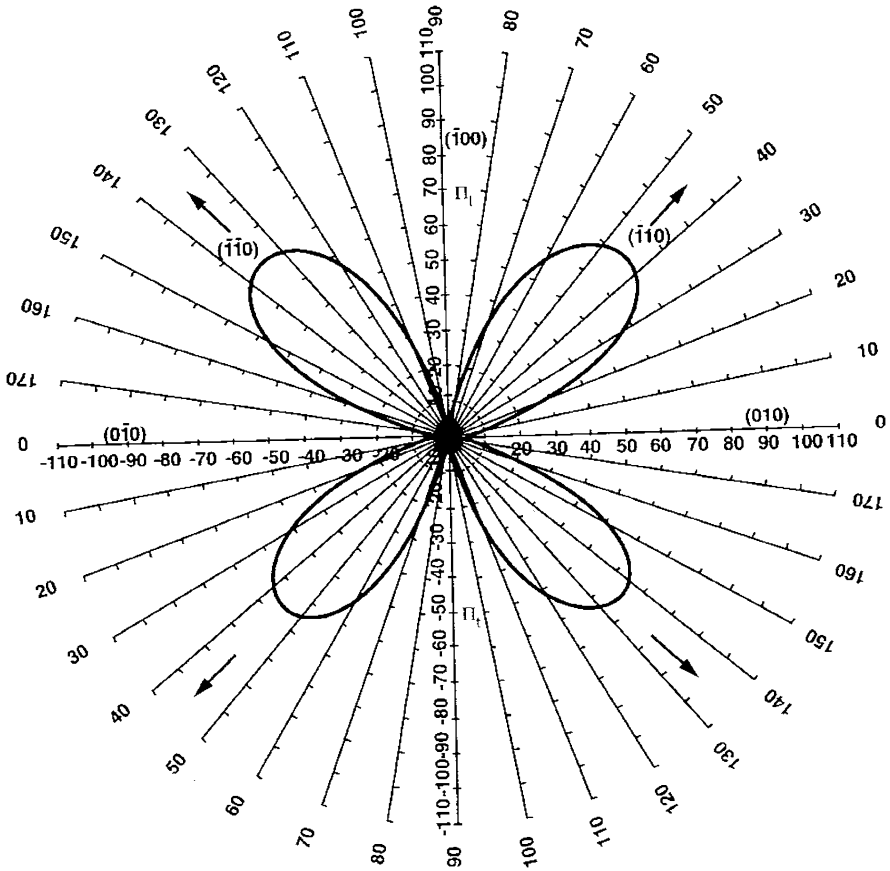
What is the weight your colleague obtains now at the terrace? Do you think that this is a good design for a microbalance?

Specific resistance of silicon at $25^\circ\text{C}$	$5 \text{ }\Omega \text{ cm}$	Poisson's ratio of polysilicon	0.23
Young's modulus of polycrystalline silicon	160 GPa	Part of the gauge factor due to change of conductivity $K_L$	30
Temperature coefficient of the resistance of polycrystalline silicon	$1.2 \times 10^{-3}$		

### Problem 7

In Fig. E10, the piezoresistive coefficients of monocrystalline silicon are shown which was cut along the (001)-plane. It is advantageous to design strain gauges such that they are orientated in [110]-direction or in a direction which is equivalent to this, because the measurement effect is large, this way. Electrical connections to the strain gauges which obviously shall show a very small sensitivity to straining may be positioned along the [100]- or an equivalent direction.

- Take from Fig. E10, the piezoresistive coefficient  $\pi_1$  for the case that it is stressed along its length.
- Find the corresponding value  $\pi_t$  for stressing perpendicular to the electrical current.



**Fig. E10** Piezoresistive coefficients in the (001)-plane for p-doped silicon [18]. The values in the upper and lower hemisphere are for longitudinal  $\pi_l$  and transversal  $\pi_t$  stress, respectively. © [1982] IEEE

(c) By a mismatch during fabrication, the strain gauges are rotated relative to the crystal orientation by  $+10^\circ$ . How are the values changed by this production error?

# Membranes

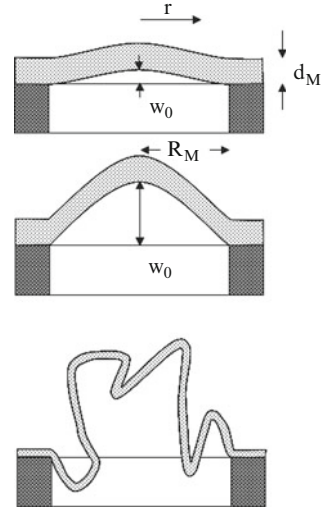
Besides conductor paths, membranes are another special type of thin films. Membranes are an important mechanical basic element in microtechnique. They are the microscopic correspondence to macroscopic gaskets, bearings, and springs. They are made of silicon, oxides, nitrides, glasses, polymers, and metals. Their thickness typically is in the range of 0.5–500  $\mu\text{m}$ . Membranes which are thinner than 0.5  $\mu\text{m}$  are very hard to manufacture without holes and are generally not strong enough to withstand usual loads. The upper limit is given by the fact that thicker membranes are no longer a microscopic element. The lateral dimensions of membranes are typically in the range between 100  $\mu\text{m}$  and 10 mm. Again the lower limit is defined by the possibilities of fabrication, while the upper limit is approximately the limit to the macroscopic world. However, all equations discussed here are valid in the macroscopic world also.

There are two types of membranes, which are frequently used in microtechnique: *Thick membranes* and *thin membranes*. A membrane is called thick when its maximum deflection  $w_0$  is much smaller than its thickness  $d_M$  and thin when the deflection is larger than the thickness. So, a thick membrane may become a thin one when the pressure drop rises and the deflection is increased. A thick membrane, sometimes, is also called a plate. The shape of the deflection of a thick membrane is determined by the bending moments acting especially at the rim where the membrane is clamped to a frame or housing. Figure 19 shows the characteristic shape of different types of membranes. As usual in microtechnique, the drawings are not to scale; small dimensions are shown larger to make them visible in the vicinity of larger ones. The deflection  $w$  of a circular thick membrane with radius  $R_M$  is described by the following equation [19]:

$$w(r) = w_0 \left(1 - \frac{r^2}{R_M^2}\right)^2. \quad (23)$$

When the membrane is deflected, its neutral fiber needs to become longer and this strain generates some stress according to Hooke's law. When the deflection becomes much larger than the thickness of the membrane, the effect on the shape of the membrane generated by the stress due to straining gets much larger than the

**Fig. 19** Thick, thin, and flexible membrane, respectively



effect of the bending moments. In this case, the bending moments may be neglected in calculations. The deflection curve of a circular thin membrane adopts the shape of a parabola and is described by the following equation:

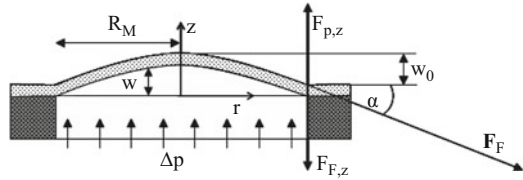
$$w(r) = w_0 \left( 1 - \frac{r^2}{R_M^2} \right). \quad (24)$$

A third membrane type is the *flexible membrane*, which is shown at the bottom of Fig. 19. It is longer than the distance between the positions where its rim is clamped and it is so thin that bending moments are of minor importance. There is no straining along the neutral fiber and no residual stress. The deflection of such a membrane is not defined. It behaves as a plastic bag which can be deformed with small forces and retains this shape until further forces act on it. A flexible membrane may be used to separate two fluids and to ensure that there is no pressure difference over the membrane. It may also be used to limit the volume of a fluid because much pressure is required when a certain displacement of the membrane shall be overcome.

The *deflection of a circular thin membrane* loaded by a constant pressure drop  $\Delta p$  can be calculated from the equilibrium of forces at the rim. The absolute value of the total force  $F_p$  acting on the membrane is the pressure difference times the membrane area  $\pi R_M^2$  (cf. Fig. 20). This force is balanced by the force  $F_F$  of the frame fixing the membrane at the circumference. The lateral components of this force cancel out when summed over the entire rim of the membrane, because there is no lateral movement of the membrane. The vertical components  $F_{F,z}$  and  $F_{p,z}$  of the force of the frame and the pressure drop, respectively, are in equilibrium:

$$F_{p,z} = \Delta p \pi R_M^2 = -F_{F,z} = -\sigma_M d_M 2 \pi R_M \sin(\alpha). \quad (25)$$

**Fig. 20** Cross-section of a circular thin membrane loaded with a pressure drop



The force of the frame in (25) is calculated from the membrane stress  $\sigma_M$  times the cross-section of the membrane around the circumference, which is the membrane thickness  $d_M$  times the length of the circumference  $2 \pi R_M$ . The vertical component of the force of the frame is obtained by multiplying with the sine of the angle at which the membrane touches the frame. For comparatively small angles  $\alpha$ , the sine is approximately the same as the tangent which is the slope of the membrane at its rim and can be calculated as the derivative of the deflection curve  $w(r)$  at the rim. Equation (25) becomes now:

$$\Delta p \pi R_M^2 \approx -\sigma_M d_M 2 \pi R_M \tan(\alpha) = -\sigma_M d_M 2 \pi R_M \left. \frac{\partial w}{\partial r} \right|_{r=R_M}. \quad (26)$$

The deflection curve of the membrane is the parabola described by (24). Calculating the derivative of this, inserting in (26) and solving for the pressure drop yields:

$$\Delta p \pi R_M^2 = -\sigma_M d_M 2 \pi R_M w_0 \frac{2}{R_M} = 4 \pi w_0 d_M \sigma_M \Rightarrow \Delta p = \frac{4 w_0 d_M}{R_M^2} \sigma_M. \quad (27)$$

The stress  $\sigma_M$  of the membrane consists of two parts, the residual stress  $\sigma_0$ , which is already present when there is no deflection of the membrane, and the stress  $\sigma_D$  due to Hooke's law generated by the deflection of the membrane. The latter part can be calculated from the strains  $\varepsilon_R$  and  $\varepsilon_T$  in radial and tangential direction, respectively, generated by the deflection. According to Hooke's law, the radial strain is calculated with the following equation:

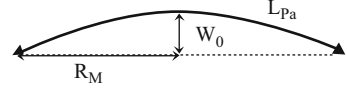
$$\varepsilon_R = \frac{\sigma_R}{E_M} + \nu_M \frac{\sigma_T}{E_M} = \frac{1}{E_M} (\sigma_R + \nu_M \sigma_T). \quad (28)$$

The first term in (28) is the strain generated by the radial stress  $\sigma_R$  of the membrane and the second term is due to the transverse strain generated by the tangential stress  $\sigma_T$  of the membrane.  $\nu_M$  and  $E_M$  denote Poisson's ratio and Young's modulus of the membrane, respectively. The tangential strain  $\varepsilon_T$  is calculated accordingly to (28):

$$\varepsilon_T = \frac{\sigma_T}{E_M} + \nu_M \frac{\sigma_R}{E_M} = \frac{1}{E_M} (\sigma_T + \nu_M \sigma_R). \quad (29)$$



**Fig. 21** Length  $L_{Pa}$  of a parabola with a peak height  $w_0$  and a base length  $2 R_M$



Radial strain is assumed to be constant over the entire membrane. This can be adequate for thin membranes only, because bending moments, which result in a strain change over the thickness of the membrane, are comparatively small and may be neglected in thin membranes. Radial membrane strain is estimated by the extension along the neutral fiber of the membrane which is necessary for deflection (cf. Fig. 21). In formularies such as [21], the length  $L_{Pa}$  of a parabola with a comparatively small height  $w_0$  at its peak and a base length  $2 R_M$  is found to be:

$$L_{Pa} \approx 2R_M \left( 1 + \frac{2}{3} \frac{w_0^2}{R_M^2} - \frac{2}{5} \frac{w_0^4}{R_M^4} \right). \quad (30)$$

Thus, the strain  $\varepsilon_R$  necessary to extend the base length to the length of the parabola is:

$$\varepsilon_R \approx \frac{2}{3} \frac{w_0^2}{R_M^2}. \quad (31)$$

The tangential strain of the membrane is currently unknown. Only two boundary conditions can be determined:

1. In the center of the membrane, radial and tangential strains are equal because the strain is not a function of direction at this position due to symmetry.
2. At the circumference, tangential strain is zero because the membrane is clamped there and the frame does not allow any movements of the membrane.

In the following, the two ansatz-equations are made that either the first or the second boundary condition would be fulfilled everywhere on the membrane. Later the results obtained this way are compared with each other.

If tangential and radial strains are equal throughout the membrane (31), (28), and (29) result in:

$$\sigma_R = \varepsilon_R \frac{E}{1 - \nu_M}. \quad (32)$$

If tangential strain is assumed to be zero everywhere, we obtain:

$$\sigma_R = \varepsilon_R \frac{E}{1 - \nu_M^2}. \quad (33)$$

For most materials, Poisson's ratio is approximately 0.3. As a consequence, the maximum change in the quantity  $(1 - \nu_M)$  versus  $(1 - \nu_M^2)$  is 11%. For polymers,

Poisson's ratio may approach 0.5 resulting in a ratio of 66%. For a general understanding of the interrelationships, this is an approximation which is still good enough.

Equation (31) is substituted in (32) and the total stress  $\sigma_M$  of the thin circular membrane is obtained from the sum of the residual stress  $\sigma_0$ , and the stress generated by the deflection of the membrane  $\sigma_R$ :

$$\sigma_M = \sigma_0 + \frac{2}{3} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M}. \quad (34)$$

This is introduced now to (27) resulting in an interrelationship between the pressure drop  $\Delta p$  over a thin, circular membrane and its radius  $R_M$ , thickness  $d_M$ , Young's modulus  $E_M$ , Poisson's ratio  $\nu_M$ , residual stress  $\sigma_0$ , and deflection  $w_0$ :

$$\Delta p = 4 \frac{w_0 d_M}{R_M^2} \left( \sigma_0 + \frac{2}{3} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M} \right). \quad (35)$$

Equation (35) is *Cabrera's equation* which is frequently used to calculate membrane deflections [22]. However, it is based on a rough approximation, and it could have been assumed also that tangential strain is zero throughout the membrane (33). As a consequence of this assumption, (35) would be:

$$\Delta p = 4 \frac{w_0 d_M}{R_M^2} \left( \sigma_0 + \frac{2}{3} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right). \quad (36)$$

The exact equation for thin, circular membranes was found by using finite element methods (FEM) [23]. An ansatz with variable parameters was made for the pressure drop  $\Delta p$  as a function of Poisson's ratio and the parameters were determined by a fit to the FEM result. The solutions for a circular and a square membrane with an edge length  $a_M$  are as follows.

Circular membrane

$$\Delta p = 4 \frac{w_0 d_M}{R_M^2} \left( \sigma_0 + \frac{2}{3} \frac{w_0^2}{R_M^2} \frac{E_M}{1.026 - 0.793\nu_M - 0.233\nu_M^2} \right). \quad (37)$$

Square membrane

$$\Delta p = 13.6 \frac{w_0 d_M}{a_M^2} \left( \sigma_0 + 1.61 \frac{w_0^2}{a_M^2} \frac{(1.446 - 0.427 \nu_M) E_M}{1 - \nu_M} \right). \quad (38)$$

It should be noted that all the calculations above are for thin membranes only, i.e., bending moments are not included. When a membrane is thick, however, bending moments dominate the behavior of the membrane and the stress may be

neglected for some cases. In the macroscopic world, the deflection of membranes is very often dominated by bending moments; therefore, solutions for this problem have existed for decades. These are the calculations for the deflection of plates, commonly found in classical mechanics books. After solving these classic macroscopic differential equations, the following equations are found [19].

Circular membrane

$$\Delta p = \frac{16}{3} \frac{d_M^3}{R_M^4} \frac{E_M}{1 - \nu_M^2} w_0 \Rightarrow w_0 = \frac{3}{16} \frac{R_M^4}{d_M^3} \frac{1 - \nu_M}{E_M} \Delta p. \quad (39)$$

Square membrane

$$\Delta p = 66 \frac{d_M^3}{a_M^4} \frac{E_M}{1 - \nu_M^2} w_0 \Rightarrow w_0 = \frac{1}{66} \frac{a_M^4}{d_M^3} \frac{1 - \nu_M^2}{E_M} \Delta p. \quad (40)$$

In general, however, both bending moments and stress affect membrane deflections. If it is desired to take both into account, the *Ritz method* can be used to find the resultant deflection [24]. In this approach, the potential energy of the membrane is calculated as a function of one or more free parameters, which are obtained by calculating the extremes of the energy function. A local minimum corresponds to a stable equilibrium of the membrane, while a maximum is an unstable one.

In the following, the Ritz method is utilized in a simple example of the deflection of a spring loaded with a mass  $m_K$  (cf. Fig. 22). The potential energy  $V_p$  of the system is the sum of the energy of the spring and the mass:

$$V_p = -m_K g_e w_0 + \frac{k}{2} w_0^2. \quad (41)$$

The derivative of the potential energy is the force. Therefore, the extremes of the potential energy correspond to the positions where the sum of the forces is zero, i.e., the equilibrium of forces. The derivative of (41) with respect to the deflection  $w_0$  of the spring is:

$$\frac{\partial V_p}{\partial w_0} = 0 = F_{Ku} + F_k = -m_K g_e + k w_0. \quad (42)$$

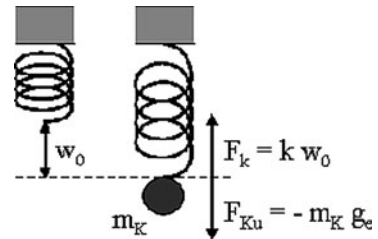


Fig. 22 Equilibrium of forces at a spring loaded with a mass

Solving this equation yields the deflection of the spring:

$$w_0 = \frac{m_K g_e}{k}. \quad (43)$$

This result could have also been found by directly calculating the equilibrium of forces, but with a more difficult problem, using the extremes of the potential energy according to the Ritz method is much easier.

The potential energy of a membrane with thickness  $d_M$ , radius  $R_M$ , Young's modulus  $E_M$ , Poisson's ratio  $\nu_M$ , and deflection  $w$  can generally be calculated from a differential equation. If membrane bending moments, residual stress, stress due to straining, and the pressure drop are included, the equation takes the following form in rectangular and polar coordinates, respectively [19]:

$$V_p = \iint dx dy \left[ \frac{E_M d_M^3}{24(1-\nu_M^2)} \left( \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 + \frac{d_M}{2} \left( \sigma_{x0} \left\{ \frac{\partial w}{\partial x} \right\}^2 + \sigma_{y0} \left\{ \frac{\partial w}{\partial y} \right\}^2 \right) \right. \\ \left. + \frac{d_M E_M}{8(1-\nu_M^2)} \left( \left\{ \frac{\partial w}{\partial x} \right\}^4 + \left\{ \frac{\partial w}{\partial y} \right\}^4 \right) - w \Delta p \right], \quad (44)$$

$$V_p = \int_0^{2\pi} d\varphi \int_0^{R_M} r dr \left[ \frac{E_M d_M^3}{24(1-\nu_M^2)} \left( \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} + \frac{1}{r^2} \frac{\partial^2 w}{\partial \varphi^2} \right)^2 \right. \\ \left. + \frac{d_M}{2} \left( \sigma_{r0} \left\{ \frac{\partial w}{\partial r} \right\}^2 + \frac{\sigma_{\varphi 0}}{r^2} \left\{ \frac{\partial w}{\partial \varphi} \right\}^2 \right) + \frac{d_M E_M}{8(1-\nu_M^2)} \left( \frac{\partial w}{\partial r} \right)^4 - w \Delta p \right]. \quad (45)$$

The first term in the brackets of the two equations above corresponds to the effect of the bending moments, the second to residual membrane stress in x- and y-, (radial and tangential) directions, respectively, the third to the contribution of the stress due to straining of the neutral fiber, and the last to the energy generated by moving the membrane at the pressure difference.

Equations (44) and (45) are differential equations. In principle, the deflection curve  $w(r)$  of the membrane would need to be found which yields the minimum potential energy. An ansatz with free parameters was used instead of the unknown deflection curve. This ansatz is chosen such that it describes the expected deflection curve as good as possible. For a circular membrane with radius  $R_M$  and center deflection  $w_0$ , an ansatz is made with a fourth-order polynomial with the free parameters  $w_0$ ,  $a_0$ ,  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$ :

$$w(r) = w_0 \left( a_0 + a_1 \frac{r}{R_M} + a_2 \frac{r^2}{R_M^2} + a_3 \frac{r^3}{R_M^3} + a_4 \frac{r^4}{R_M^4} \right). \quad (46)$$

Some of the free parameters can be derived from boundary conditions: The slope of the membrane at its center needs to be zero, i.e. the derivative of  $w(r)$  at  $r = 0$

needs to be zero; therefore  $a_1 = 0$ . Also, the deflection at the center of the membrane is defined to be  $w_0$ , resulting in  $a_0 = 1$ . Finally, the deflection at the rim and the slope of the membrane at the rim are necessarily zero. Using these parameters, the following equations are created:

$$a_2 = a_4 - 3 \quad \text{and} \quad a_3 = 2 - 2a_4. \quad (47)$$

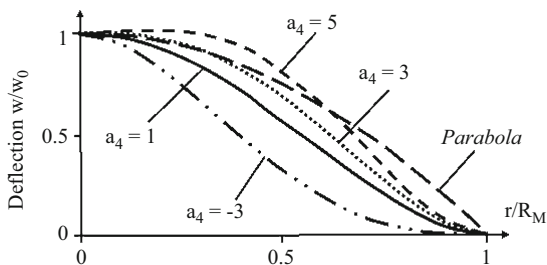
As a consequence,  $w_0$  and  $a_4$  are the only free parameters remaining, because  $a_2$  and  $a_3$  depend on  $a_4$ , and (46) becomes:

$$w(r) = w_0 \left( 1 + [a_4 - 3] \frac{r^2}{R_M^2} + [2 - 2a_4] \frac{r^3}{R_M^3} + a_4 \frac{r^4}{R_M^4} \right), \quad (48)$$

where  $w_0$  is the center deflection of the membrane and  $a_4$  is a measure of the deflection shape of the membrane. The case when  $a_4 = 1$  corresponds to a circular, thick membrane [cf. (23) on page 29]. Figure 23 shows the deflection curve as a function of the parameter  $a_4$ . A thin membrane [(24) on page 30] would correspond to the parabola shown in the figure. Obviously, a parabola cannot be perfectly modeled by this ansatz, because the slope of a parabola is not zero at the rim. However, the parabola is best approximated by the ansatz when  $a_4$  is between 3 and 5.

The ansatz (48) is now differentiated and the results are inserted into (45). The integral can then be calculated because the resulting function is simply a polynomial. This calculation comprises a lot of simple steps which all bear the risk of mistakes. Therefore, computer codes are used to perform the analytical calculations. The following equation is the result of such a code:

$$\begin{aligned} V_p = 2\pi & \left[ \frac{E_M d_M^3}{24(1-\nu_M^2)} \frac{w_0^3}{R_M^2} \left( 9 + \frac{6}{5}a_4 + \frac{7}{15}a_4^2 \right) + \frac{d_M}{2} \sigma_{r0} w_0^2 \left( \frac{3}{5} + \frac{2}{35}a_4 + \frac{1}{105}a_4^2 \right) \right. \\ & + \frac{E_M d_M}{1-\nu_M^2} \frac{w_0^4}{R_M^2} \left( \frac{9}{70} + \frac{6}{385}a_4 + \frac{3}{385}a_4^2 + \frac{2}{5,005}a_4^3 + \frac{1}{30,030}a_4^4 \right) \\ & \left. - \Delta p w_0 R_M^2 \left( \frac{3}{20} + \frac{1}{60}a_4 \right) \right]. \quad (49) \end{aligned}$$



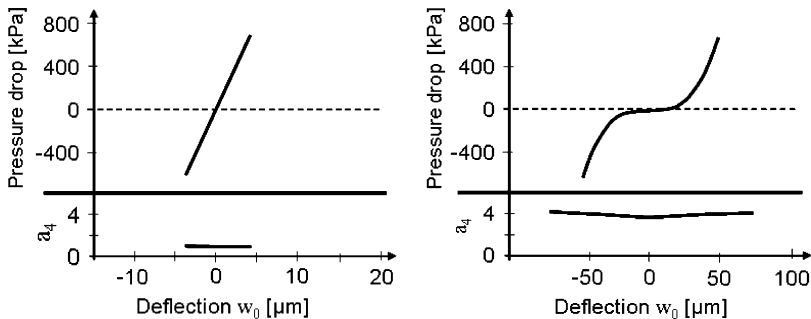
**Fig. 23** Deflection curve of a circular membrane calculated with (48) as a function of the parameter  $a_4$

The equation above describes the potential energy of the membrane as a function of the free parameters  $w_0$  and  $a_4$ . As before, we are looking for those values of the parameters for which the potential energy is minimal. As usual, the minimum of a function is found by calculating the zero of the derivative with respect to the corresponding parameter, i.e. the following system of two equations needs to be solved for  $w_0$  and  $a_4$ :

$$0 = \frac{\partial V_p}{\partial w_0} = 2 \pi \left[ \frac{E_M d_M^3}{24(1-\nu_M^2)} \frac{w_0}{R_M^2} \left( 9 + \frac{6}{5} a_4 + \frac{7}{15} a_4^2 \right) + \frac{d_M}{2} \sigma_{r0} w_0 \left( \frac{3}{5} + \frac{2}{35} a_4 + \frac{1}{105} a_4^2 \right) + \frac{E_M d_M}{1-\nu_M^2} \frac{w_0^3}{R_M^2} 4 \left( \frac{9}{70} + \frac{6}{385} a_4 + \frac{3}{385} a_4^2 + \frac{2}{5,005} a_4^3 + \frac{1}{30,030} a_4^4 \right) - \Delta p R_M^2 \left( \frac{3}{20} + \frac{1}{60} a_4 \right) \right], \quad (50)$$

$$0 = \frac{\partial V_p}{\partial a_4} = 2 \pi \left[ \frac{E_M d_M^3}{24(1-\nu_M^2)} \frac{w_0^2}{R_M^2} \left( \frac{6}{5} + \frac{14}{15} a_4 \right) + d_M \sigma_{r0} w_0^2 \left( \frac{1}{35} a_4 + \frac{1}{105} a_4 \right) + \frac{E_M d_M}{1-\nu_M^2} \frac{w_0^4}{R_M^2} \left( \frac{6}{385} + \frac{6}{385} a_4 + \frac{6}{5,005} a_4^2 + \frac{2}{15,015} a_4^3 \right) - \frac{\Delta p w_0 R_M^2}{60} \right]. \quad (51)$$

Unfortunately, this system of equations is too complicated to be solved analytically, but numerically it can be solved easily. In Fig. 24, the interrelationship between pressure drop and both membrane deflection  $w_0$  and the parameter  $a_4$  are shown for two typical cases. On the left, there is the linear relationship between pressure drop and deflection of a 20- $\mu\text{m}$  thick silicon membrane with a radius of 500  $\mu\text{m}$ , no residual stress, and a ratio of Young's modulus  $E_M$  to  $(1-\nu_M^2)$  of 240 GPa.  $a_4$  is approximately one for all pressure values, as expected for a thick membrane.



**Fig. 24** Interrelationship between the pressure drop over a circular membrane and its center deflection  $w_0$  and the parameter  $a_4$  describing the deflection curve of the membrane both calculated by numerically solving the system of equations (50) and (51). On the left, the result is shown for a thick silicon membrane and on the right for a thin polyimide membrane

On the right of Fig. 24 are the results for a 5- $\mu\text{m}$  thick polyimide membrane also with a radius of 500  $\mu\text{m}$  but with a tensile residual stress of 50 MPa and a ratio of Young's modulus  $E_M$  to  $(1-\nu_M^2)$  of 3.3 GPa. The interrelationship between pressure drop and deflection is not linear in this case. This is in agreement with Cabrera's equation for a thin membrane [(35) on page 33], which contains a nonlinear term due to the straining of the neutral fiber. Accordingly the value of the parameter  $a_4$  is around 4, which approximates a parabola shaped deflection curve of the membrane (cf. Fig. 23).

As the system of (50) and (51) cannot be solved analytically, a simpler ansatz is made calculating the membrane deflection with the Ritz method. It is assumed that the parameter  $a_4$  always equals one. As noted above, this is a good approximation for a thick membrane but does not match well for the case of a thin membrane. Therefore, precision is lost with a more general solution which can be used to study more general cases. The only free parameter remaining in the ansatz is the deflection  $w_0$  of the membrane. At the extremes of the potential energy, the derivative of the potential energy with respect to the deflection is zero:

$$0 = \frac{\partial V_p}{\partial w_0} = \frac{2}{3} \pi \left( \frac{8 E_M d_M^3 w_0}{3 (1 - \nu_M^2) R_M^2} + 2 d_M \sigma_0 w_0 + \frac{128 E_M d_M w_0^3}{105 (1 - \nu_M^2) R_M^2} - \frac{1}{2} \Delta p R_M^2 \right) = F_M. \quad (52)$$

The derivative of the potential energy with respect to the movement is the force, in general. Therefore, (52) is a description of the force  $F_M$  acting on the membrane, and the deflection where the force becomes zero corresponds to the equilibrium of forces as in the simple case of (42) on page 34.

The four terms in the brackets of (52) correspond to the contributions of bending moments, residual stress  $\sigma_0$ , stress due to straining of the neutral fiber, and pressure drop  $\Delta p$  over the membrane, respectively.

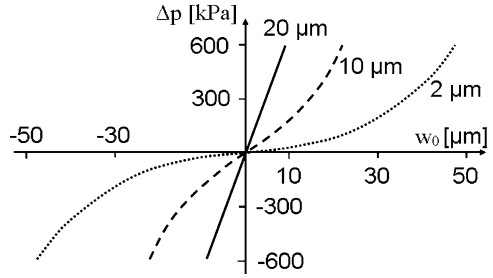
Solving (52) for the pressure drop  $\Delta p$  over the membrane yields an equation which describes the relationship between pressure drop and the deflection of a membrane:

$$\Delta p = \frac{4 d_M w_0}{R_M^2} \left( \frac{4 d_M^2}{3 R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64 w_0^2}{105 R_M^2} \frac{E_M}{1 - \nu_M^2} \right). \quad (53)$$

This equation can be used to investigate the behavior of membranes. The three terms in the brackets describe the contribution of the bending moments, residual stress  $\sigma_0$ , and stress due to straining of the neutral fiber, respectively. Bending moments and residual stress show a linear interrelationship between pressure drop and membrane deflection, while the stress due to straining contributes with its third power. That is, if the first and the second term of the bracket of (53) are much larger than the third term, the pressure is a linear function of the deflection.

This is important, if a pressure sensor has to be designed which shows a linear characteristic curve. The characteristic curve (53) gets more linear by increasing the

**Fig. 25** Pressure drop over a circular membrane as a function of its deflection and different thicknesses calculated with (53)



thickness  $d_M$  of the membrane, which enlarges the first term in the brackets compared with the others. In Fig. 25, the pressure drop is shown as a function of the deflection for different thicknesses of a membrane as calculated with (53). Residual stress  $\sigma_0$ , radius  $R_M$ , and ratio of Young’s modulus  $E_M$  to  $(1-\nu_M)$  of this membrane are 300 MPa, 600  $\mu\text{m}$ , and 120 GPa, respectively.

As expected, the figure shows that the pressure drop is a more linear function of the deflection for thicker membranes and that this function is linear approximately up to a deflection as large as the thickness of the membrane. So it is an option to design a thicker membrane, if a linear interrelationship between pressure drop and deflection is desired. However, Fig. 25 also shows that the deflection of thicker membranes is less at a certain pressure drop. That is, if a thick membrane is used for a pressure sensor, it is less sensitive. Therefore, there is a compromise between linearity and sensitivity.

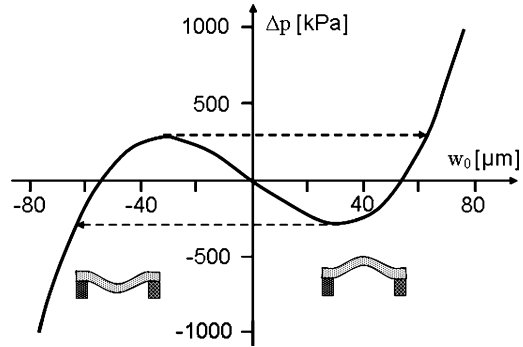
The same trend is observed when (53) is investigated with respect to changes in residual stress. A larger residual stress results in a more linear function as well as less sensitivity for a pressure sensor equipped with such a membrane.

If a *compressive residual stress* is assumed for the membrane, negative values need to be used for  $\sigma_0$  in (53). This is again a rough approximation because it is assumed that the compressive stress of the membrane would be constant throughout the membrane. In general, this is not true, but the negative input value of  $\sigma_0$  may be considered as a kind of effective residual stress which generates a similar effect as the true distribution of the stress. So a precise description of membrane deflection cannot be expected, but the general trend will be shown.

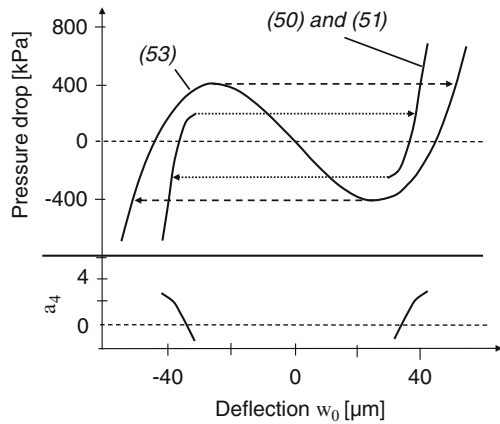
Figure 26 shows the change which occurs when a compressive residual stress of  $\sigma_0 = -600$  MPa is used instead of the tensile stress assumed for Fig. 25. In a certain pressure range, there are two stable states for the membrane: It may buckle up (positive deflection, right side of the graph) or buckle down (negative deflection, left side of the graph). This is called a *bistable membrane*. When there is no pressure drop over the membrane and it is buckling down, it will deflect approximately -55  $\mu\text{m}$ . If the pressure is then raised, the downward deflection becomes increasingly smaller until at approximately -30  $\mu\text{m}$  the membrane snaps over to the opposite side, following the dashed arrow and arriving at an upward deflection of approximately 60  $\mu\text{m}$ . Starting from that point, the pressure can be reversed to about -300 kPa before the membrane snaps back to a downward deflection.



**Fig. 26** Pressure drop over a circular membrane as a function of its deflection calculated with (42) as in Fig. 21 but with a compressive residual stress of  $-600$  MPa



**Fig. 27** Interrelationship between the pressure drop over a circular membrane with compressive stress and its center deflection as calculated with (53) and (50) and (51) on page 37



The curve between the two points where the membrane is snapping over cannot be reached by adjusting the pressure difference over the membrane to a certain value. However, it is a real part of the function and can be reached by holding the membrane at a certain deflection by some other means. The pressure drop shown in the graph corresponds to the force per unit area of the membrane necessary to hold the membrane in position and prevent further deflection. The closer the membrane is brought to the non-deflected position, the smaller is the force necessary to hold it at that position. The non-deflected position corresponds to an unstable equilibrium of forces of the membrane.

The curves shown in Fig. 26 were also calculated by a numerical solution of (50) and 51 (see page 37) with the same compressive residual stress. In Fig. 27, these curves are compared with each other. The numerical solution is a better approximation because it includes the change of the deflection shape of the membrane. Snapping over is calculated to occur already at a larger deflection and at a smaller pressure difference than calculated with the simpler (53). The calculation also shows that the deflection shape changes from parabolic shape to something beyond the shape of a thick membrane when the snapping over position is approached.

**Fig. 28** Interrelationship between the pressure drop over a circular membrane and its center deflection for different residual stresses as calculated with (53)

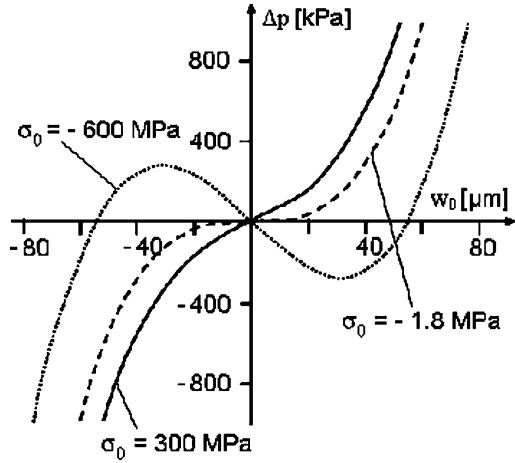


Figure 28 shows the interrelationship between the pressure drop over a circular membrane and its center deflection for different residual stresses. At a residual stress of  $-1.8$  MPa, there is a transition from a monostable to a bistable membrane. This stress, called the critical stress  $\sigma_k$  of the membrane, occurs where the slope of the function at the origin is zero. Therefore, the critical stress at which the transition takes place can be calculated from the zero of the derivative of (53):

$$0 = \left. \frac{\partial \Delta p}{\partial w_0} \right|_{w_0=0} = \frac{4 d_M}{R_M^2} \left( \frac{4 d_M^2}{3 R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 \right) \Rightarrow \sigma_k = -\frac{4 d_M^2}{3 R_M^2} \frac{E_M}{1 - \nu_M^2}. \quad (54)$$

When the residual stress is smaller (more compressive) than the critical stress, the membrane becomes bistable and buckles if there is no pressure drop. A membrane with larger (more tensile) stress is flat and does not deflect without a pressure drop. Note that the critical stress is as large as the contribution of the bending moments in (53). In other words, buckling takes place when the residual stress overcomes the bending moments of the membrane.

The deflection  $w_0(\Delta p = 0)$  of a membrane without pressure drop is calculated now by inserting a pressure drop equal to zero into (53):

$$\Delta p = 0 = \frac{4 d_M w_0}{R_M^2} \left( \frac{4 d_M^2}{3 R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64 w_0^2}{105 R_M^2} \frac{E_M}{1 - \nu_M^2} \right). \quad (55)$$

There are two possibilities to solve this equation:

$$w_0 = 0 \quad (56)$$

and

$$\frac{4}{3} \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64}{105} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} = 0$$

$$\Rightarrow w_0 = \pm \frac{\sqrt{35}}{4} d_M \sqrt{\frac{\sigma_0}{-\frac{4}{3} (d_M^2/R_M^2) (E_M/(1 - \nu_M^2))} - 1}. \quad (57)$$

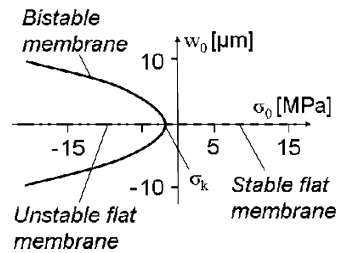
The first solution corresponds to a flat membrane, while the second describes the deflection of a bistable membrane with compressive stress. The denominator under the square root is the critical stress (cf. 54), and this solution gives a suitable result only if the residual stress is more compressive than the critical stress (cf. Fig. 29).

Equation (56) is a valid solution for all residual stresses of the membrane. In fact, if the residual stress is more tensile than the critical stress, it is the only solution, because the term under the root is negative. If the residual stress is more compressive than the critical one, the membrane is in an unstable, non-deflected position.

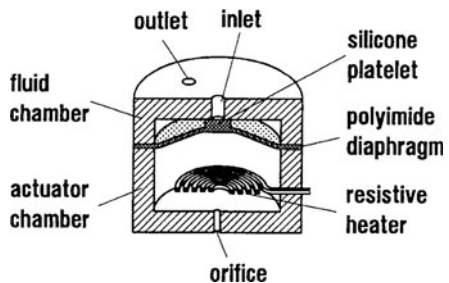
The residual stress of a membrane is changed as a function of the temperature, if the frame or housing it is fixed at shows a different thermal expansion. Equation (57) may be useful to calculate the temperature at which the membrane buckles.

The buckling of a bistable membrane with compressive stress was employed to design the microvalve shown in Fig. 30 [25]. The compressive stress of the membrane results in a force which presses the membrane against the orifice of the inlet and keeps it closed against an outer pressure. When the valve is opened, the heater in the actuator chamber is switched on and the air escapes through the orifice to the environment. After this, the heater is switched off and the decreasing pressure in the actuator chamber pulls the membrane into the downward stable position. For closing, the membrane is pushed up again by activating the heater. The electrical

**Fig. 29** Deflection of a circular membrane without pressure difference as a function of its residual stress calculated with (57) and (56) (dashed line)



**Fig. 30** Microvalve equipped with a bistable membrane [25]



current through the heater is then lowered slowly, allowing the pressure in the actuator chamber to equalize through the orifice and avoiding the pulling down of the membrane.

To design the membrane of a valve of this type its deflection and elastic force need to be calculated. The optimum distance between the undeflected membrane and the inlet orifice with respect to a maximized force on the inlet also needs to be known.

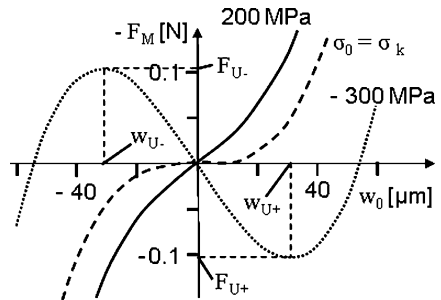
If a force such as the pressure from the inlet of the valve is acting on the center of a circular membrane, the deflection can be approximated with (52) on page 38. Fig. 31 shows the calculated elastic force generated by a 2- $\mu\text{m}$  thick, circular membrane (radius of 600  $\mu\text{m}$  and a ratio  $E_M/(1 - \nu_M^2)$  of 120 GPa) as a function of membrane deflection  $w_0$  and stress  $\sigma_0$ , if there is no pressure difference over the membrane. The situation is similar to the case of a membrane loaded with a pressure shown in Fig. 28. The only difference is that the membrane is deflected by a force instead of a pressure. When the stress is more compressive than the critical stress, it becomes bistable. If there is no force acting on the membrane and it is buckling downward (negative  $w_0$ ), its deflection is approximately  $-55 \mu\text{m}$ . If it is deflected upward, then it generates a negative elastic force which tries to deflect it down further. The absolute value of this force is increasing until the deflection  $w_{U-}$  is reached at a force  $F_{U-}$  and the membrane snaps over to the upper side.

The extremes of the deflection are found by calculating the derivative of (52) on page 38 and finding the zeros:

$$\frac{\partial F_M}{\partial w_0} = \frac{2\pi}{3} \left( \frac{8 E_M d_M^3}{3 (1 - \nu_M^2) R_M^2} + 2 d_M \sigma_0 + \frac{128 E_M d_M w_0^2}{35 (1 - \nu_M^2) R_M^2} \right) = 0. \quad (58)$$

$$\begin{aligned} \Rightarrow w_U &= \pm \frac{\sqrt{35}}{\sqrt{3} \cdot 4} d_M \sqrt{\frac{\sigma_0}{-\frac{4}{3} (d_M^2/R_M^2) (E_M/1 - \nu_M^2)} - 1} \\ &= \pm \frac{\sqrt{35}}{\sqrt{3} \cdot 4} d_M \sqrt{\frac{\sigma_0}{\sigma_k} - 1} = \frac{1}{\sqrt{3}} w_0 (\Delta p = 0). \end{aligned} \quad (59)$$

The comparison of the above result with (57) shows that the deflection at the snapping over point is just a factor  $1/\sqrt{3} \approx 0.58$ , smaller than the deflection when



**Fig. 31** Elastic force of a circular membrane as a function of its deflection and stress calculated with (52) on page 38

no force or pressure difference is acting on the membrane. The maximum force generated by the membrane is achieved at  $w_U$  and is found by inserting (59) into (52):

$$F_U = \pm \frac{\pi}{3} \left( \frac{7}{3} \sqrt{\frac{35}{3}} d_M^2 \sigma_k \left( \frac{\sigma_0}{\sigma_k} - 1 \right)^{3/2} - \Delta p R_M^2 \right). \tag{60}$$

Figure 32 shows the effect of a pressure difference on snapping over and maximum force of a membrane with a residual stress of  $-300$  MPa as calculated with (52).

The pressure drop simply adds a constant to (52). Thus, the deflections  $w_{U+}$  and  $w_{U-}$  at which the membrane snaps over are not changed. The maximum forces  $F_{U+}$  and  $F_{U-}$ , however, become asymmetric and the bistable membrane positions without an outer force acting are changed also.

If the stress of a membrane becomes increasingly compressive, more complex bulging may occur which is not described with a simple ansatz such as (48) on page 36. The membrane then contains several folds as seen in Fig. 33, which shows the membrane of a micropump heated and thus being exposed to large compressive stress [26].

When a membrane is designed for a pressure sensor, a microvalve, or a micropump, it is important to know at which *burst pressure* the membrane breaks. The

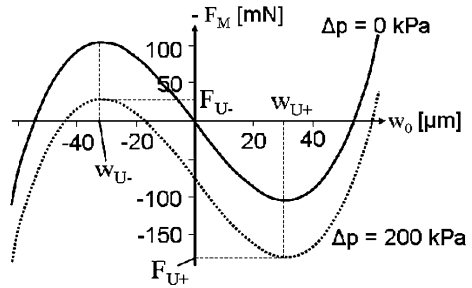


Fig. 32 Elastic force of a circular membrane as a function of its deflection and stress calculated with (52) on page 38

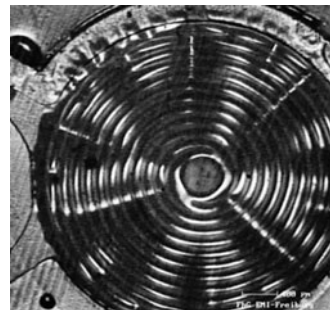


Fig. 33 Membrane of a micropump showing several folds due to large compressive stress generated by heating of the membrane [26]. © [1997] IEEE

burst pressure can be calculated from Cabrera's equation or another equation describing the relationship between deflection  $w_0$ , pressure drop  $\Delta p$ , and total stress of the membrane  $\sigma_M$ . Here the calculation based on Cabrera's equation [(35) on page 33] assumes that the membrane is a thin membrane. This assumption will be suitable for most applications because the deflection of a membrane will be larger than its thickness when it bursts. However, membranes from brittle materials such as silicon dioxide and ceramics may be exceptions.

The membrane bursts when the total stress  $\sigma_M$  becomes larger than the yield stress  $\sigma_y$ . [The total stress is described by (34) on page 33.] At the burst pressure  $\Delta p_y$ , the membrane will have the burst deflection  $w_{0y}$  and (34) and (35) become:

$$\sigma_y = \sigma_0 + \frac{2}{3} \frac{w_{0y}^2}{R_M^2} \frac{E_M}{1 - \nu_M}. \quad (61)$$

$$\Delta p_y = 4 \frac{w_{0y} d_M}{R_M^2} \sigma_y. \quad (62)$$

The burst deflection is calculated from (61)

$$w_{0y} = R_M \sqrt{\frac{3}{2} (\sigma_0 - \sigma_y) \frac{1 - \nu_M}{E_M}} \quad (63)$$

and inserted into (62), resulting in an expression with which the burst pressure can be calculated as follows:

$$\Delta p_y = 4 \frac{d_M}{R_M} \sigma_y \sqrt{\frac{3}{2} (\sigma_0 - \sigma_y) \frac{1 - \nu_M}{E_M}}. \quad (64)$$

To calculate the burst pressure, the yield stress of the membrane material needs to be known. The yield stress can be determined by increasing the pressure differential in small increments and measuring the deflection. When the membrane breaks, the last deflection and pressure measured are the burst deflection and burst pressure, respectively, and the yield stress can be calculated by solving (62).

The calculations shown above for a circular membrane can also be performed for square membranes with edge length  $a_M$ . A parabolic deflection is assumed in  $x$ - and  $y$ -direction:

$$w(x, y) = w_0 \left(1 - 4 \frac{x^2}{a_M^2}\right)^2 \left(1 - 4 \frac{y^2}{a_M^2}\right)^2. \quad (65)$$

The potential energy is calculated from (44) (page 35) with this ansatz and differentiated with respect to the central deflection  $w_0$  to obtain the force acting on the membrane. It is assumed that the stress  $\sigma_x$  and  $\sigma_y$  of the square membrane may be different in the  $x$ - and  $y$ -direction, respectively:

$$\begin{aligned}
\frac{\partial V_p}{\partial w_0} = F_M &= \frac{64}{75} \left( \frac{2^{10}}{49} \frac{E_M}{1-v_M^2} \frac{d_M^3 w_0}{a_M^2} + \frac{2^{10}}{49 \times 9} d_M w_0 (\sigma_x + \sigma_y) \right. \\
&\quad \left. + \frac{2^{28}}{7 \times 9 \times 11^2 \times 13^2 \times 17} \frac{d_M E_M w_0^3}{1-v_M^2 a_M^2} - \frac{1}{3} \Delta p a_M^2 \right) \\
&= 17.8 \frac{E_M d_M^3 w_0}{1-v_M^2 a_M^2} + 1.98 d_M (\sigma_x + \sigma_y) w_0 + 10.46 \frac{E_M d_M w_0^3}{1-v_M^2 a_M^2} - 0.284 \Delta p a_M^2.
\end{aligned} \tag{66}$$

If the force acting on the membrane is zero, i.e. the membrane is in equilibrium, this equation can be solved for  $\Delta p$ :

$$\Delta p = 3 \frac{d_M}{a_M^2} w_0 \left( 20.9 \frac{E_M}{1-v_M^2} \frac{d_M^2}{a_M^2} + 2.32 (\sigma_x + \sigma_y) + 12.3 \frac{E_M}{1-v_M^2} \frac{w_0^2}{a_M^2} \right). \tag{67}$$

From this equation, the critical stress of a square membrane can be calculated. Similar to the case of a circular membrane (cf. page 41), the critical stress  $\sigma_k$  is achieved when the second term in the parenthesis of (67) is more compressive than the first one. In other words, the average of the stress components in the x- and y-directions is more negative than the following:

$$\sigma_k = -\frac{9}{2} \frac{E_M}{1-v_M^2} \frac{d_M^2}{a_M^2}. \tag{68}$$

The deflection  $w_0$  ( $\Delta p = 0$ ) of a square membrane with a compressive stress larger than the critical stress  $\sigma_k$  and no pressure difference, the deflection  $w_U$  at the snapping over point, and the force  $F_U$  needed for snapping over are calculated in the same way as for a circular membrane:

$$\begin{aligned}
w_0(\Delta p = 0) &= \pm \frac{11 \times 13}{3 \times 2^8} \sqrt{\frac{17}{7}} d_M \sqrt{\frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1} \\
&= \pm 0.29 d_M \sqrt{\frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1}.
\end{aligned} \tag{69}$$

$$\begin{aligned}
w_U &= \pm \frac{11 \times 13}{3 \times 2^8} \sqrt{\frac{17}{3 \times 7}} d_M \sqrt{\frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1} = \pm 0.17 d_M \sqrt{\frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1} \\
&= \frac{1}{\sqrt{3}} w_0(\Delta p = 0).
\end{aligned} \tag{70}$$

$$\begin{aligned}
F_U &= \pm \frac{64}{9 \times 25} \left( \frac{8 \times 11 \times 13 \times 23}{3^5 \times 49} \sqrt{\frac{17}{21}} d_M^2 \sigma_k \left( \frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1 \right)^{3/2} - \Delta p a_M^2 \right) \\
&= \pm 0.28 \left( 1.99 d_M^2 \sigma_k \left( \frac{\sigma_x + \sigma_y}{2 \sigma_k} - 1 \right)^{3/2} - \Delta p a_M^2 \right).
\end{aligned} \tag{71}$$

**Fig. 34** Membrane with a boss at its center

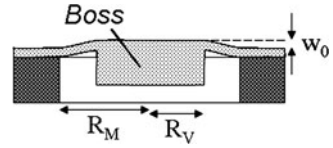


Figure 25 and the last term of (53) (page 38) show that, in general, the deflection of a membrane is not a linear function of the pressure drop. For several applications, especially pressure sensors, it is desirable to design membranes with a deflection which is a linear function of the pressure difference. Therefore, several designs were developed providing improved linearity.

One possibility is to fabricate a *membrane with a boss* at its center (cf. Fig. 34). The boss stiffens the membrane so much that it may be assumed that only the annular part of the membrane not covered by the boss is strained when a pressure difference is applied to the membrane. As a result, the coefficients in (53) on page 38 need to be altered [27]:

$$\Delta p = \frac{d_M w_0}{R_M^2} \left( a_p \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + 4 \sigma_0 + b_p \frac{w_0^2}{R_M^2} E_M \right). \tag{72}$$

With

$$a_p = \frac{16}{3} \frac{1}{1 - (R_V^4/R_M^4) - 4(R_V^2/R_M^2) \ln(R_M/R_V)} \tag{73}$$

and

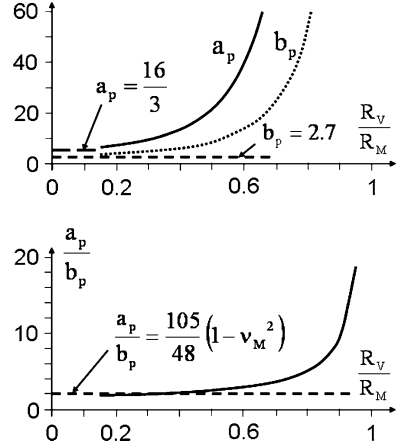
$$b_p = \frac{(7 - \nu_M)(1 + (R_V^2/R_M^2) + (R_V^4/R_M^4))/3 + (3 - \nu_M)^2/(1 + \nu_M)(R_V^2/R_M^2)}{(1 - \nu_M)(1 - (R_V^4/R_M^4))(1 - (R_V^2/R_M^2))}. \tag{74}$$

In the above equations,  $R_V$  denotes the radius of the boss. Equation (72) is an approximation which is valid when the radius of the boss  $R_V$  is larger than 15% of the radius of the membrane  $R_M$  and the thickness of the boss is at least six times larger than the thickness of the membrane.

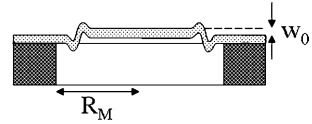
The effect of the boss is illustrated in Fig. 35, which shows the coefficients  $a_p$ ,  $b_p$ , and their ratio as a function of the ratio  $R_V/R_M$  of the radii of boss and membrane. Poisson’s ratio was assumed to be 0.3 for this example calculation. The coefficients are much larger with a sufficiently large boss at the membrane center than without it. ( $16/3$  and  $256/(105(1 - \nu_M^2)) \approx 2.7$ , respectively, as shown in (53) on page 38.) This means that the deflection calculated with (72) will be smaller with a boss, resulting in less sensitivity if the membrane is employed in a pressure sensor. On the other hand, the residual stress of the membrane has a comparatively smaller effect on the deflection. This is an advantage because the cross-sensitivity of the sensor to stress changes is reduced. Such stress changes often occur when the



**Fig. 35** Coefficients  $a_p$ ,  $b_p$ , and their ratio calculated with (73) and (74)



**Fig. 36** Membrane with corrugations at its rim



housing of the sensor shows a different thermal expansion than the membrane or when the sensor is mounted somewhere.

The ratio of the coefficients  $a_p$  and  $b_p$  increases with the radius of the boss. As a consequence, the characteristic curve (membrane deflection as a function of the pressure difference) becomes more linear, because  $a_p$  is the coefficient of a linear term and  $b_p$  of a nonlinear term in (72). However, a significant effect is achieved, only if the boss covers more than approximately 80% of the membrane (cf. Fig. 35).

Another interesting way to achieve a more linear membrane deflection as a function of the pressure difference is to design a *corrugated membrane* (cf. Fig. 36). Similar to the case of a membrane with a boss, (53) on page 38 needs to be altered to take into account the effect of the corrugations [28]:

$$\Delta p = \frac{d_M w_0}{R_M^2} \left( a_p \frac{d_M^2}{R_M^2} E_M + 4\sigma_0 + b_p \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right). \quad (75)$$

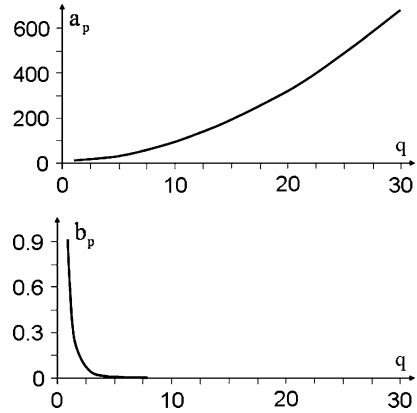
With

$$a_p = \frac{2}{3} \frac{(q+3)(q+1)}{1 - (\nu_M^2/q^2)} \quad (76)$$

and

$$b_p = \frac{165(q+1)(q+3)}{q^2(q+4)(q+11)(2q+1)(3q+5)}. \quad (77)$$

**Fig. 37** Coefficients  $a_p$ ,  $b_p$ , and their ratio calculated with (76) and (77)



In these equations,  $q$  is the profile factor of the corrugations in the membrane. It is a measure of the size of the corrugations. The calculation is described in [27]. A membrane without any corrugations corresponds to  $q = 1$ . Typical corrugated membranes correspond to a profile factor between 10 and 30.

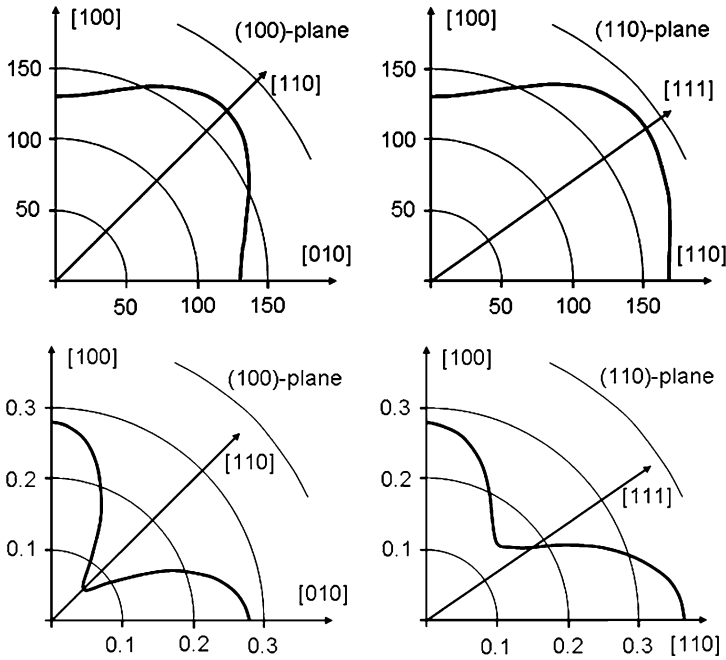
Figure 37 shows the coefficients  $a_p$  and  $b_p$  as a function of  $q$ . Only small corrugations are required to reduce the nonlinear third term in the parenthesis of (75) and to allow dominance of the linear first term. Thus, only a few shallow corrugations at the rim of a membrane are enough to achieve a linear displacement as a function of the pressure difference. Since  $a_p$  rises quickly with increasing  $q$ , the corrugated membrane becomes much stiffer and the effect of residual stress is diminished significantly.

It should be noted that in all calculations described in this chapter, it has been assumed that Young’s modulus and Poisson’s ratio are isotropic properties of the membrane material. This is true for amorphous and polycrystalline materials. However, for monocrystalline materials such as silicon, which is the material most widely used in microtechnique, these properties are a strong function of crystal orientation. Figure 38 shows Young’s modulus and Poisson’s ratio as a function of crystal orientation for monocrystalline silicon [29].

Different orientations result in Young’s moduli which vary up to 40% and Poisson’s ratios can change by as much as a factor 4. Fortunately, in most equations, the ratios  $E/(1 - \nu)$  or  $E/(1 - \nu^2)$  occur, which only vary by 23% in the worst case because Young’s modulus is largest for the [110] direction where Poisson’s ratio is smallest, and the ratio balances these variations slightly.

It is very hard to include the anisotropic properties in analytical calculations. Therefore, effective or average values are often used instead.

The following table provides an overview on the equations which can be used to calculate the interrelationship between the pressure difference over a membrane and its deflection.



**Fig. 38** Young's modulus [GPa] (*upper row*) and Poisson's ratio (*lower row*) of monocrystalline silicon as a function of crystal orientation. Reprinted with permission from [29]. Copyright [1965], American Institute of Physics

**Table 2** Equations for the calculation of the interrelationship of pressure difference  $\Delta p$  and deflection  $w_0$  of a membrane

Membrane type	Equation
General solution for circular membrane, rough approximation	$\Delta p = 4 \frac{d_M w_0}{R_M^2} \left( \frac{4}{3} \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64}{105} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right)$
General solution for square membrane, rough approximation	$\Delta p = 3 \frac{d_M w_0}{a_M^2} \left( 20.9 \frac{E_M}{1 - \nu_M^2} \frac{d_M^2}{a_M^2} + 2.32(\sigma_x + \sigma_y) + 12.3 \frac{E_M}{1 - \nu_M^2} \frac{w_0^2}{a_M^2} \right)$
Thin, circular, without bending moments, exact solution	$\Delta p = 4 \frac{d_M w_0}{R_M^2} \left( \sigma_0 + \frac{2}{3} \frac{w_0^2}{R_M^2} \frac{E_M}{1.026 - 0.793\nu_M - 0.233\nu_M^2} \right)$
Thin, square, without bending moments, exact solution	$\Delta p = 13.6 \frac{d_M w_0}{a_M^2} \left( \sigma_0 + 1.61 \frac{w_0^2}{a_M^2} \frac{(1.446 - 0.427\nu_M)E_M}{1 - \nu_M} \right)$
Thick, circular, without stress, exact solution	$\Delta p = \frac{16}{3} \frac{d_M^3}{R_M^4} \frac{E_M}{1 - \nu_M^2} w_0 \Rightarrow w_0 = \frac{3}{16} \frac{R_M^4}{d_M^3} \frac{1 - \nu_M^2}{E_M} \Delta p$
Thick, square, without stress, exact solution	$\Delta p = 66 \frac{d_M^3}{a_M^4} \frac{E_M}{1 - \nu_M^2} w_0 \Rightarrow w_0 = \frac{1}{66} \frac{a_M^4}{d_M^3} \frac{1 - \nu_M^2}{E_M} \Delta p$
Circular membrane, with boss, for $a_p$ and $b_p$ see (73) and (74)	$\Delta p = \frac{d_M w_0}{R_M^2} \left( a_p \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + 4 \sigma_0 + b_p \frac{w_0^2}{R_M^2} E_M \right)$
Circular membrane, with corrugations, $[a_p, b_p]$ see (76) and (77)]	$\Delta p = \frac{d_M w_0}{R_M^2} \left( a_p \frac{d_M^2}{R_M^2} E_M + 4 \sigma_0 + b_p \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right)$

## Exercises

### Problem 8

You are asked to design a pressure sensor which fulfills the following specifications:

Measurement range	0–100 kPa
Sensitivity	1 kPa
Maximum deviation of the characteristic curve from linearity	1% of the value measured
Temperature range for the use	–20 to 50°C

Due to the available equipment in your company, you decide to design a circular silicon membrane. The deflection of this membrane shall be ruled by the bending moments. The following material properties are given:

Young's modulus of the membrane	150 GPa
Poisson's ratio	0.3
Thermal strain of the membrane	$3 \times 10^{-6}/^{\circ}\text{C}$
Thermal strain of the housing	$5 \times 10^{-6}/^{\circ}\text{C}$

- Assume that the sensor membrane is strained directly by the difference in thermal strain of sensor and housing, because it is much thinner than the housing. Which change of the stress of the membrane do you expect over the temperature range specified for the sensor?
- The membrane may be deflected up to what portion of the thickness in order to keep the linearity specified? (Neglect the initial stress of the membrane.)
- How small may the ratio of thickness to radius of the membrane become to avoid that the thermal strain changes the sensor signal for less than 1%.
- How small may the ratio of thickness to radius of the membrane become to avoid that the maximum deflection of the membrane [from (b)] is not exceeded at the upper limit of the measurement range? (Again neglect the initial stress of the membrane.)
- What is the smallest deflection of the membrane which needs to be measured to achieve the desired sensitivity? (The diameter of the membrane is 2 mm.)
- What thickness is needed for the membrane?

### Problem 9

In the lecture, you got to know a membrane with bistable states.

- What is the reason for the snapping over of the membrane?

- (b) The snapping over can be used to design a bistable valve (cf. Fig. 30 on page 42). In such a bistable valve [30], the membrane consists of polyimide and shows a thickness of  $d_M = 25 \mu\text{m}$ . The radius of the membrane is  $R_M = 1.5 \text{ mm}$ .
- (c) Please calculate the critical stress  $\sigma_K$  of this membrane.
- (d) What is the deflection of the membrane at the snapping over if the initial stress is  $\sigma_0 = -8.5 \text{ MPa}$ ?

---

Young's modulus of polyimide	1.66 GPa	Poisson's ratio of polyimide	0.41
------------------------------	----------	------------------------------	------

---

# Strain Gauges on Membranes

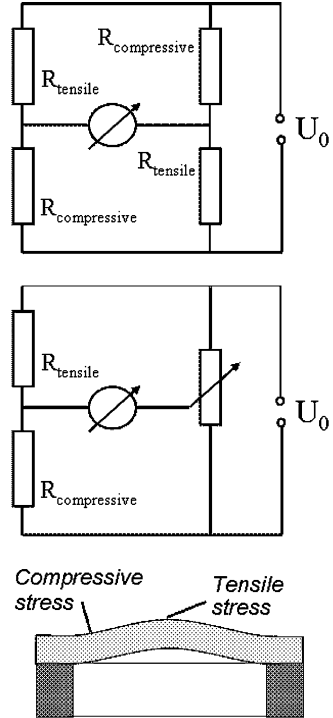
The deflection of membranes is most often measured by strain gauges on or in the membrane. The application used most frequently is the deflection of membranes in silicon pressure sensors. In pressure sensors, deflections of the membrane of less than a micrometer need to be detected and the strain generated by these deflections is on the order of  $10^{-4}$ . As a consequence, the resistance of conductor paths that are employed as strain gauges may change much more due to temperature changes than due to strain. Therefore, the deflection of such a membrane cannot be measured reliably without temperature compensation. The usual way for temperature compensation is to build up a Wheatstone bridge from two or four strain gauges as shown in Fig. 39.

The Wheatstone bridge contains strain gauges which are arranged differently such that their resistances change differently when the membrane is deflected. For example, two strain gauges are placed at a position on the membrane where there is compressive stress and two other strain gauges where there is tensile stress (cf. Fig. 39). Thus, a voltage difference is generated between the two branches of the bridge when the membrane is deflected. If the temperature is changed, all strain gauges undergo the same changes resulting in no change of the output voltage. The bridge may be driven either with a constant voltage or with a constant current source.

The resistances are designed and fabricated such that there is no output voltage when the membrane is not deflected. As a consequence, the voltage change can be measured starting from a level of 0 V. If the change of a single resistance due to membrane deflection would be measured, a small change of a much larger voltage had to be measured which is much more difficult to detect.

The four strain gauges shown in the upper part of Fig. 39 are called a full bridge, while the two strain gauges in the middle part of the figure are a half bridge. At a half bridge, only the strain gauges are situated on the membrane and the potentiometer and other electronics are not on the membrane. The advantage of a half bridge is that small deviations of the resistances which occur during fabrication can be adapted with the potentiometer easily and that more area is available to achieve a sufficiently large resistance with a metal conductor path. The resistance of the strain gauges needs to be large enough, because power consumption and even heating of the membrane by the current through the strain gauges could become too large.

**Fig. 39** Wheatstone bridges; *Top*: Full bridge; *Middle*: Half bridge; *Bottom*: Positions on a deflected thick membrane with compressive and tensile stress



On the other hand, the output of a full bridge is two times larger than for a half bridge, and modern fabrication techniques allow balancing the bridge during the fabrication process. Nearly all membranes for pressure sensors are made of silicon nowadays and the resistance of a strain gauge from silicon is large enough even if it is short. That is why most pressure sensors are equipped with a full bridge.

The output voltage of a Wheatstone bridge which shall be used to measure the deflection of a membrane is largest, if the difference in resistance change of the strain gauges is maximized. Therefore, the designer of strain gauges needs to search for positions and orientations on the membrane which show the most different resistance changes for strain gauges.

In the following, it is assumed that the strain gauges show no mechanical effect on the straining of the membrane, although this is not quite correct in any case. It is assumed that the strain gauges are so thin and weak that they do not hinder the deflection of the membrane and that their lateral strain is the same as the strain of the membrane without a strain gauge at that position.

The change of a resistance due to the geometrical effect of *metal strain gauges* is described by (16) on page 21. The resistance change  $\Delta R_{e1}$  of a strain gauge with Poisson's ratio  $\nu$  on a membrane with lateral strain  $\varepsilon_l$  and transversal strain  $\varepsilon_t$  is composed of terms which describe the effect of the strain gauge to become longer

and wider by the straining and by getting thinner perpendicular to each of these strain directions:

$$\Delta R_{el} = R_{el}(\varepsilon_l + \nu\varepsilon_l - \varepsilon_t + \nu\varepsilon_t) = R_{el}(\varepsilon_l(1 + \nu) - \varepsilon_t(1 - \nu)). \quad (78)$$

Note that this result is different from (17) on page 22, which describes the resistance change of a strained wire not fixed on a substrate.

Equation (78) shows that straining longitudinal to the direction of the electrical current in a conductor path results in an enhancement of the electrical resistance, while straining perpendicular to the strain gauge reduces the resistance (a bit less, because of the negative sign of  $\nu$ ). So, at a first glance, arranging strain gauges parallel and perpendicular to the direction of strain will result in a large difference of the resistance change when the membrane is deflected. However, the strain distribution on the membrane needs to be considered also.

Until now stress and strain on the surface of thick membranes have not been discussed in detail. This was sufficient for the approximate calculation of the deflection and form with (53) on page 38. If the proper arrangement of strain gauges on such membranes shall be calculated, more exact equations are needed. They can be calculated with the Ritz method and are found in text books, e.g., [19]. Radial  $\sigma_R$  and tangential  $\sigma_T$  stress of a circular, thick membrane without residual stress and with radius  $R_M$ , thickness  $d_M$ , and Poisson's ratio  $\nu_M$  as a function of the pressure difference  $\Delta p$  over the membrane are:

$$\sigma_R = \frac{3}{8} \frac{\Delta p}{d_M^2} (R_M^2(1 + \nu_M) - r^2(3 + \nu_M)) \quad (79)$$

$$\sigma_T = \frac{3}{8} \frac{\Delta p}{d_M^2} (R_M^2(1 + \nu_M) - r^2(1 + 3\nu_M)). \quad (80)$$

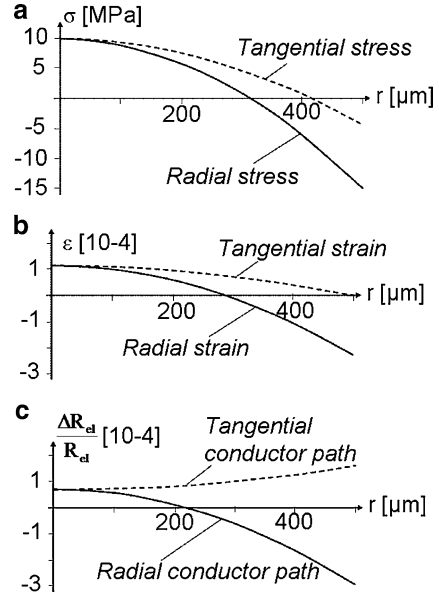
Figure 40a shows the radial and tangential stress of a circular thick membrane as calculated with (79) and (80). A glass membrane was assumed for the calculation with a Young's modulus and Poisson's ratio of 60 GPa and 0.3, respectively, and a radius and thickness of 500 and 50  $\mu\text{m}$ , respectively, with no residual stress. As expected radial and tangential stress are the same at the center of the membrane, because of the symmetry. At the center the surface of the membrane shows tensile stress, because it is bulged convexly, while at the rim there is compressive stress in radial direction due to concave bowing. There is no tangential strain at the rim, because the membrane is fixed there. As a consequence, a tangential stress appears which compensates the transverse strain which would occur due to the radial strain.

The strain on the membrane can be calculated from these equations now by Hooke's law which has been employed already on page 31 where Cabrera's equation was derived [cf. (28) and (29)]:

$$\varepsilon_R = \frac{\sigma_R}{E_M} - \nu_M \frac{\sigma_T}{E_M} = \frac{3}{8} \frac{\Delta p}{E_M d_M^2} (1 - \nu_M^2) (R_M^2 - 3r^2) \quad (81)$$



**Fig. 40** (a) Stress and (b) strain on the upper surface of a circular thick membrane and (c) relative resistance change of strain gauges on that membrane



$$\varepsilon_T = \frac{\sigma_T}{E_M} - \nu_M \frac{\sigma_R}{E_M} = \frac{3}{8} \frac{\Delta p}{E_M d_M^2} (1 - \nu_M^2) (R_M^2 - r^2). \quad (82)$$

Figure 40b shows the strain components of the glass membrane. As expected there is no tangential strain at the rim. The radial strain on the upper surface of the membrane is tensile and compressive at the center and the rim, respectively, because of convex and concave bulging, respectively.

The strains shown in (81) and (82) are inserted into (78) now both for conductor paths in radial and in tangential direction on the membrane. For a radial conductor path, the radial strain is taken as the longitudinal strain and the tangential strain is taken for the transverse strain. For a tangential conductor path, the tangential strain is longitudinal and the radial one is transversal. This results in the following equations for the change  $\Delta R_{el,R}$  and  $\Delta R_{el,T}$  of the electrical resistance  $R_{el}$  of radial and transversal strain gauges on the surface of thick circular membranes:

$$\Delta R_{el,R} = \frac{3}{8} R_{el,R} \frac{\Delta p}{E_M d_M^2} (1 - \nu_M^2) (R_M^2 \nu - (1 + 2\nu)r^2) \quad (83)$$

$$\Delta R_{el,T} = \frac{3}{8} R_{el,T} \frac{\Delta p}{E_M d_M^2} (1 - \nu_M^2) (R_M^2 \nu + (1 - 2\nu)r^2). \quad (84)$$

The relative resistance change calculated with these equations is shown in Fig. 40c for the glass membrane. The resistance of tangential strain gauges is rising

everywhere on the membrane when it is bulged by a pressure difference. This effect increases by approximately a factor of 2 from the center to the rim of the membrane. Radial strain gauges in the center of the membrane gain an increase of their resistance at the center and a decrease at the rim. The largest resistance change is the decrease of a radial strain gauge at the rim.

Equations (79)–(84) describe stress and strain on a thick circular membrane and the resistance change of strain gauges on such a membrane as a function of the pressure drop  $\Delta p$ . Actually these quantities are a function of membrane deflection  $w_0$  and the deflection is generated by pressure drop. Therefore, it appears to be expedient to write these equations also as a function of deflection. This is easily achieved by replacing  $\Delta p$  by (53) (page 38). The second and third term in the parenthesis of (53) have to be ignored because the calculation of (79)–(84) are for a circular thick membrane without residual stress:

$$\sigma_R = 2 \frac{d_M w_0}{R_M^4} \frac{E_M}{1 - \nu_M^2} (R_M^2(1 + \nu_M) - r^2(3 + \nu_M)), \quad (85)$$

$$\sigma_T = 2 \frac{d_M w_0}{R_M^4} \frac{E_M}{1 - \nu_M^2} (R_M^2(1 + \nu_M) - r^2(1 + 3\nu_M)), \quad (86)$$

$$\varepsilon_R = 2 \frac{d_M w_0}{R_M^4} (R_M^2 - 3r^2), \quad (87)$$

$$\varepsilon_T = 2 \frac{d_M w_0}{R_M^4} (R_M^2 - r^2), \quad (88)$$

$$\Delta R_{el,R} = 2 R_{el,R} \frac{d_M w_0}{R_M^4} (R_M^2 \nu - (1 + 2\nu)r^2), \quad (89)$$

$$\Delta R_{el,T} = 2 R_{el,T} \frac{d_M w_0}{R_M^4} (R_M^2 \nu + (1 - 2\nu)r^2). \quad (90)$$

As it has to be expected, strain and resistance change of strain gauges are no function of Young's modulus and Poisson's ratio of the membrane material. Only the stress is a function of the elastic properties of the membrane.

As mentioned above, two strain gauges need to be arranged on the membrane such, that their resistance changes differ as much as possible producing the largest possible output voltage of the Wheatstone bridge. According to Fig. 40c, it appears to be advantageous to place both radial and tangential strain gauges near the rim of the membrane, because their resistance changes are most different there.

However, it needs to be considered also that a metal strain gauge should have a sufficiently large resistance. The bridge must be driven with a comparatively large voltage providing a suitable output. This would result in a large electrical current through the strain gauge, if its resistance is not large enough. The current needs to

be small, because it generates heating of the membrane resulting in a larger deflection. The heating may also be asymmetrically over the membrane and cause a change of the output voltage. Therefore, the strain gauges must be comparatively long and there may be not enough area on a small membrane to place both radial and tangential conductor paths at the rim. This can also be a reason, why a half bridge with two strain gauges on the membrane is more suitable than a full bridge with four strain gauges.

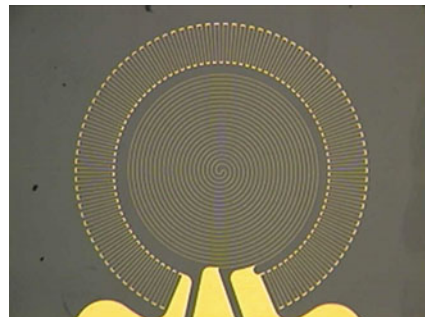
If there is not enough area to place both radial and tangential strain gauges at the rim of the membrane, it is best to design radial strain gauges at the rim and tangential ones next to them towards the membrane center. Such an arrangement is shown in Fig. 41 where a spiral (which is nearly a tangential conductor path) covers the central area of the membrane and radial strain gauges are placed at the rim.

The strain gauges are designed such that their resistances show nearly the same value, because this way the output of the bridge becomes maximized. In the light of maximum production tolerances, it is better to place tangential strain gauges at the rim and radial ones nearer to the center, because the variation of the resistance change with the radial position is less, this way. Thus, a shift of the strain gauges relative to the membrane center results in a smaller change of the characteristic curve which describes the output voltage as a function of membrane deflection.

The designer needs to take into account that the tangential connections between radial conductor paths contribute to the total effect also. Therefore, these tangential connections had been designed wider than the radial strain gauges in Fig. 41. It needs to be considered also that the calculations shown here are approximate, because the Young's modulus of the strain gauges reduces membrane deflection, and, therefore, the real resistance changes are smaller than calculated here.

On thin membranes, the stress and strain distribution is different than in the case of the thick membranes described by (85), (86) and (87), (88), respectively. The strain is the same on both sides of a thin membrane, and there is no position with compressive stress or strain. It is assumed now that the radial strain is constant over the entire membrane and that it can be calculated from the increase in length which is necessary to deflect a membrane with parabolic form as calculated with (31) on page 32.

**Fig. 41** Strain gauges from gold on a membrane (The membrane cannot be recognized). The width of the strain gauges is  $5\ \mu\text{m}$  and the diameter of the membrane is 2 mm. (Courtesy of Karlsruhe Institute of Technology, KIT)



$$\epsilon_R \approx \frac{2}{3} \frac{w_0^2}{R_M^2}. \tag{31}$$

The tangential strain in the center of the membrane needs to be the same as the radial one and it needs to be zero at the rim, because the membrane is fixed there. It is assumed now that the tangential strain diminishes from the center to the rim nearly as a parabolic function yielding:

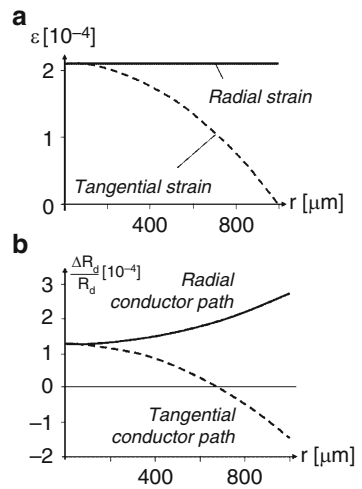
$$\epsilon_T = \epsilon_R \left( 1 - \frac{r^2}{R_M^2} \right) = \frac{2}{3} \frac{w_0^2}{R_M^2} \left( 1 - \frac{r^2}{R_M^2} \right). \tag{91}$$

In the same way as above for a thick membrane, the strains shown in (31) and (91) are inserted into (78) now both for conductor paths in radial and in tangential direction on the membrane. This results in the following equations:

$$\Delta R_{el,R} = \frac{2}{3} R_{el,R} \frac{w_0^2}{R_M^2} \left( 2\nu + (1 - \nu) \frac{r^2}{R_M^2} \right), \tag{92}$$

$$\Delta R_{el,T} = \frac{2}{3} R_{el,T} \frac{w_0^2}{R_M^2} \left( 2\nu - (1 + \nu) \frac{r^2}{R_M^2} \right). \tag{93}$$

Strain and resistance changes of strain gauges on a membrane with a radius  $R_M$  of 1 mm and a thickness  $d_M$  of 10  $\mu\text{m}$  for a deflection of 17.7  $\mu\text{m}$  are shown in Fig. 42 as a function of the radial position as calculated by (31), (91) and (92), (93), respectively. As in the case of the thick membrane, the resistance changes are largest at the rim and show opposite signs for radial and tangential conductor paths, respectively. If there is not sufficient area to place strain gauges with both



**Fig. 42** (a) Strain and (b) relative resistance change of strain gauges on a thin circular membrane

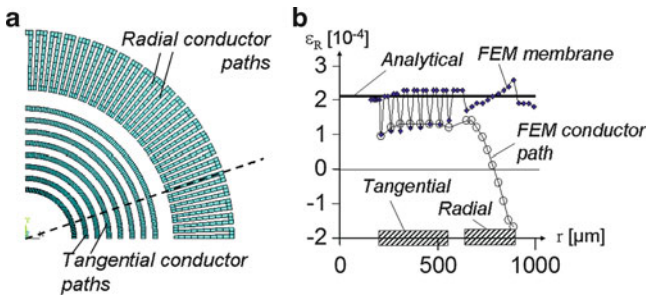
orientations at the rim, because otherwise their resistance would be too small, it is preferable to arrange radial strain gauges at the rim and tangential ones next to them towards the center, because the difference of the resistance changes is a little bit larger this way.

Instead of at the surface of a membrane, strain gauges can also be arranged in its neutral fiber at half of the membrane thickness. This way, the strain at the surface generated by membrane bending is not transferred to the strain gauges and only the strain due to the lateral expansion of the membrane during deflection is measured by the strain gauges. As a consequence, the resistance change has to be calculated with (92) and (93) also for deflections much smaller as membrane thickness.

Equations (92) and (93) are proportional to the square of membrane deflection. As a consequence, the resistance change is not a function of the direction in which the membrane is deflected.

The remarks made above for thick membranes on the contribution of tangential connections and the equal resistances of radial and tangential strain gauges are true for thin membranes also. The mechanical effect of the conductor paths on membrane deflection is obviously even more important for thin membranes. Figure 43 shows (a) the model and (b) the radial strain as calculated by finite element methods (FEM) of a circular thin membrane and radial and tangential conductor paths on it. The conductors are from gold, 200 nm thick and 20  $\mu\text{m}$  wide, and the membrane is made of polyimide, 10  $\mu\text{m}$  thick, has a diameter of 2 mm, and is loaded with a pressure difference of 1 kPa [20].

The radial strain at the middle of the membrane and the conductor path are shown in Fig. 43b. The assumption of a constant radial strain (31) which was made for the analytical calculation is compared with the FEM results in Fig. 43b. It is clearly seen that the assumption was not too bad; the strain in the middle of the membrane is nearly constant. However, where there is a strain gauge on top of the membrane, the strain is reduced for the tangential conductor paths and changed for the radial ones, because their Young's modulus hinders straining. This means that the real strain which contributes to the measurement is much smaller than the effect calculated analytically for a thin membrane.



**Fig. 43** (a) Layout and (b) strain of a membrane and strain gauges on it along the *dashed line* shown in (a) as calculated by FEM [20]

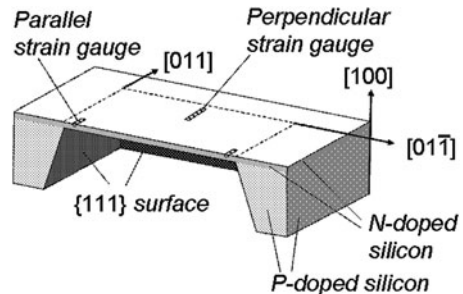
At the position of the radial strain gauges at the rim, the strain of the membrane is tensile, while the strain of the conductor path is compressive. This means that bending of the membrane occurs which had been assumed to be negligible.

Until now, the geometrical effect of strain gauges has been discussed. Nearly all pressure sensors are made of monocrystalline silicon, however. Their geometrical effect is negligible compared with the change of the specific resistance due to straining which is a function of the orientation relative to the crystal (cf. Fig. 17 on page 25).

*Membranes from monocrystalline silicon* are fabricated most easily and precisely by the so-called electro-chemical etch stop. For this process, the  $\{100\}$ -surface layer is n-doped, while the rest of the silicon wafer is p-doped. A voltage is supplied between the surface layer and a solution which under these conditions is etching the p-doped silicon only. As a result of this process, an n-doped square membrane remains with edges orientated in  $[110]$  or a physically equivalent direction (cf. Fig. 44). So, for a pressure sensor, the deflection of an n-doped membrane needs to be measured.

As in the case of circular membranes, the deflection of a thick square membrane is a linear function of the pressure load, because bending moments are dominating [cf. (53) on page 38 and the following text]. Tangential strain is zero at the rim and radial strain is maximal at the rim and at the center.

Strain gauges are introduced into the membrane by p-doping of the n-doped membrane. Therefore, the piezoresistive effect of p-doped monocrystalline silicon needs to be considered. The piezoresistive constants of p-doped silicon on an  $\{100\}$ -plane are shown in Fig. 17 on page 25. The effect is largest in  $[110]$ -direction which is parallel to the edges of the square membrane. Because radial strain is maximal and tangential strain is zero at the rim of the membrane, it is best to place two radial (perpendicular to the edge) and two tangential (parallel to the edge) strain gauges on the square membrane. This way, the strain gauges perpendicular to the edge are strained only longitudinally and the strain gauges parallel to the edge are strained only transversally. The piezoresistive effect of longitudinally and transversally strained conductor paths in  $[110]$ -direction in p-doped silicon shows opposite sign. Thus, they can be combined to a full bridge as shown in Fig. 39 and explained on page 53.



**Fig. 44** Typical silicon pressure sensor with a p-doped square membrane etched out of a  $(100)$ -wafer and n-doped strain gauges in  $(110)$ -direction parallel and perpendicular to the edges

## Exercises

### Problem 10

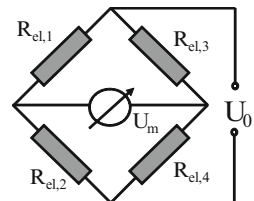
The relative change of the resistance of strain gauges is small. To measure them certain circuits are used. The most common of them is the Wheatstone bridge (Fig. E11).

The so-called bridge of the circuit shown as the connection of the resistances  $R_{el,1}$ ,  $R_{el,2}$ ,  $R_{el,3}$ , and  $R_{el,4}$  is supported on one diagonal by the constant voltage  $U_0$ . The voltage measured on the other diagonal is the output signal  $U_m$  which is a measure for the difference of the resistances  $R_{el,1}$  to  $R_{el,4}$ . The bridge is in balance, if the output voltage  $U_m$  is zero. If the resistances  $R_{el,1}$  to  $R_{el,4}$  are changed after the equalization of the bridge,  $U_m$  is proportional to  $\Delta R_{el,i}$ .

- Use Ohm's law to derive an equation for the electrical current through the branches of the bridge  $I_{el,1}$  (current through  $R_{el,1}$  and  $R_{el,2}$ ) and  $I_{el,2}$  (current through  $R_{el,3}$  and  $R_{el,4}$ ). Hint: At equilibrium  $U_m = 0$ .
- Use the law of Kirchhoff ( $\sum U_n = I_{el,n} R_{el,n}$ ) to derive the output voltage  $U_m$ . Insert your result from (a) into the equation of  $U_m$ . Which condition needs to be fulfilled by the resistances  $R_{el,1}$ ,  $R_{el,2}$ ,  $R_{el,3}$ , and  $R_{el,4}$  to equalize the bridge?
- Show that the output of a bridge is maximum, if all resistances  $R_{el,1}$ ,  $R_{el,2}$ ,  $R_{el,3}$ , and  $R_{el,4}$  are the same and the relative changes  $\Delta R_{el,1}/R_{el,1}$ ,  $\Delta R_{el,2}/R_{el,2}$ ,  $\Delta R_{el,3}/R_{el,3}$ , and  $\Delta R_{el,4}/R_{el,4}$  are constant.
- Assume that all resistances of the Wheatstone bridge are the same ( $R_{el,1} = R_{el,2} = R_{el,3} = R_{el,4} = R_{el}$ ). Then the resistances are changed by  $\Delta R_{el,1}$ ,  $\Delta R_{el,2}$ ,  $\Delta R_{el,3}$ , and  $\Delta R_{el,4}$ . Which equation is now valid for the change of the output voltage  $\Delta U_m$ ? Hint: Use the equation for  $U_m$  which you derived in (b) and insert the resistance changes  $\Delta R_{el,1}$ ,  $\Delta R_{el,2}$ ,  $\Delta R_{el,3}$ , and  $\Delta R_{el,4}$ .
- Assume that the resistances  $R_{el,1}$  and  $R_{el,4}$  are enhanced by the same amount  $\Delta R_{el}$  by which the resistances  $R_{el,2}$  and  $R_{el,3}$  are reduced:

$$\Delta R_{el,1} = \Delta R_{el,4} = \Delta R_{el} \text{ and } \Delta R_{el,2} = \Delta R_{el,3} = -\Delta R_{el}.$$

How is the expression derived in (c) for the output voltage  $\Delta U_m$  changed by this assumption?



**Fig. E11** Electronic circuit of a Wheatstone bridge

- (f) What is the advantage of the bridge circuit compared with the measurement of the resistance change of a single strain gauge?

**Problem 11**

In a pressure sensor four strain gauges ( $\Delta R_{el,1}$ ,  $\Delta R_{el,2}$ ,  $\Delta R_{el,3}$ , and  $\Delta R_{el,4}$ ) are arranged at the rim of a circular (100)-silicon membrane (cf. Fig. E12). These strain gauges are connected such that they form a Wheatstone bridge. The strain gauges are fabricated by p-doping of a silicon membrane and orientated parallel to the  $\langle 110 \rangle$ -direction.

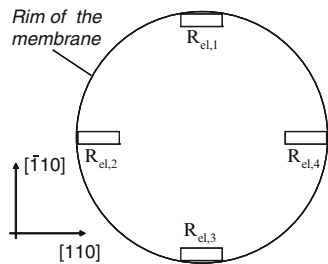
Piezoresistive effects occur parallel and perpendicular to the conductor path. The relative change of the resistance of a single strain gauge  $\Delta R_{el}/R_{el}$  is described by:

$$\frac{\Delta R_{el}}{R_{el}} = \pi_l \sigma_l + \pi_t \sigma_t,$$

where  $\pi_l$  and  $\pi_t$  denote the piezoresistive coefficients parallel and perpendicular to the conductor path, respectively, while  $\sigma_l$  and  $\sigma_t$  denote the stress parallel and perpendicular to the conductor path, respectively. As the conductor path is orientated parallel to the  $\langle 110 \rangle$ -direction, the piezoresistive coefficients are  $\pi_l = -66 \times 10^{-11} \text{ Pa}^{-1}$  and  $\pi_t = 72 \times 10^{-11} \text{ Pa}^{-1}$  (cf. Table 1 on page 25).

- (a) Derive expressions both for the resistance changes  $\Delta R_{el,1}/R_{el,1}$  and  $\Delta R_{el,3}/R_{el,3}$  and for the resistance changes  $\Delta R_{el,2}/R_{el,2}$  and  $\Delta R_{el,4}/R_{el,4}$ .  
 Take into account that radial  $\sigma_R$  and tangential  $\sigma_T$  stress arises when the circular membrane is loaded with a pressure difference. The equations for the radial  $\sigma_R$  and tangential  $\sigma_T$  stresses of a circular membrane can be found in (79) and (80), respectively, (page 55). Assume that the distance of the strain gauges to the center would be the radius of the membrane  $R_M$ .
- (b) As an approximation, it shall be assumed now that the absolute value of the piezoresistive coefficients have an average value  $\pi_m$ :

$$\pi_l \approx -\pi_t = \pi_m = 69 \times 10^{-11} \text{ Pa}^{-1}.$$



**Fig. E12** Arrangement of p-doped conductor paths on a circular silicon membrane



With this approximation, the following interrelationships between the resistance changes apply:

$$\frac{\Delta R_{el,1}}{R_{el,1}} = \frac{\Delta R_{el,3}}{R_{el,3}} \approx -\frac{\Delta R_{el,2}}{R_{el,2}} = -\frac{\Delta R_{el,4}}{R_{el,4}} = \frac{\Delta R_{el}}{R_{el}}.$$

Based on your results from (a) derive an equation which describes the characteristic curve of the pressure sensor. This is the output voltage as a function of the pressure difference over the membrane.

- (c) Calculate, using the result of (b), the change of the output voltage  $\Delta U_m$  at a pressure difference of 150 kPa.

The voltage support is  $U_0 = 10$  V. The radius of the silicon membrane is 400  $\mu\text{m}$  and its thickness is 25  $\mu\text{m}$ . Poisson's ratio of silicon is 0.23.

# Beams

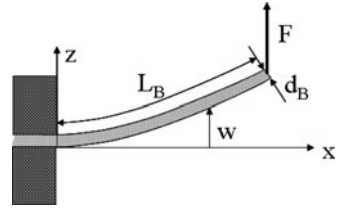
Beams are an important mechanical basic element in microtechnique. They are the microscopic correspondence to macroscopic bearings and springs. They can be made of nearly every rigid material such as silicon, oxides, nitrides, glasses, polymers, and metals. Their thickness typically is in the range of 1–500  $\mu\text{m}$ . Beams which are thinner than 1  $\mu\text{m}$  are very hard to manufacture and are not strong enough, in general. The upper limit is given by the fact that thicker beams are no longer a microscopic element. The width of beams is not much smaller than their thickness but may become a factor of 100 larger than it. When it is bent transversally, the dimension of the beam in the direction of bending is called its thickness and the dimensions perpendicular to this are called width and length. The length of beams typically is in the range between 10  $\mu\text{m}$  and 20 mm. Again the lower limit is defined by the possibilities of fabrication, while the upper limit is approximately the limit to the macroscopic world. However, all equations discussed here are valid in the macroscopic world also.

As beams are used as bearings, it is important to know what kinds of elastic deformations are possible and how they can be calculated in advance. The deflection of a beam loaded with a force is a function of the bearing which holds it in place. Text books of classical mechanical engineering such as [32–34] contain a variety of bearings and kinds of loadings on beams and show the corresponding functions which describe the interrelationship between load and deflection. In microtechnique, clamped beams are important only, because there is no way to simply support or fix a beam. On the other hand, the equations found in textbooks of macroscopic mechanics usually include the effect of bending moments only and do not take into account residual stress of beams clamped at both ends which is important in microcomponents.

There are three ways of *deflecting a beam*: longitudinal, transversal (in  $x$  and  $y$ ), and torsional. The deflection applied most is the transversal one shown schematically in Fig. 45. The deflection  $w$  of a beam with length  $L_B$ , width  $b_B$ , thickness  $d_B$ , and Young's modulus  $E_B$  as a function of the force  $F$  acting at its free end is given by the following equation:

$$w = \frac{F}{6 E_B I} x^2 (3 L_B - x). \quad (94)$$

**Fig. 45** Schematic drawing of a transversally deflected beam



In this equation, the force is due to bending of the beam, i.e., straining in the direction of the beam. As a consequence, if the beam is made of monocrystalline material where Young's modulus is not isotropic (cf. Fig. 38 on page 50), the value in longitudinal direction of the beam has to be used in the equation.

In (94),  $x$  is the coordinate in the direction of the non-deflected beam and  $I$  is the area momentum of inertia as calculated with (8) on page 7. The area momentum of inertia contains the information on the cross-sectional shape of the beam.

In the case of a symmetrical cross-section, the neutral fiber is just at the middle of the beam. Beams etched from silicon often have the shape of a trapezoid with a longer width  $d_{B,1}$  and a shorter width  $d_{B,2}$ . In this case, the distance of the neutral fiber  $d_{B,n}$  from the shorter width is:

$$d_{B,n} = \frac{d_B}{3} \frac{2 d_{B,1} + d_{B,2}}{d_{B,1} + d_{B,2}}. \quad (95)$$

The area momentums of inertia of beams with the cross-sections used most often in microtechnique are shown in Table 3.

More equations for area momentums of inertia are found in textbooks such as [32, 33].

Very often the deflection  $w_0$  of the free end of a beam clamped at the other end needs to be calculated as a function of the force  $F$  acting at the free end. It can easily be obtained by inserting the length  $L_B$  of the beam for  $x$  into (94):

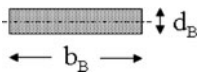
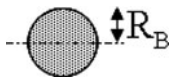
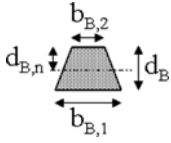
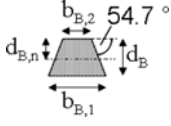
$$w_0 = \frac{L_B^3 F}{3 E_B I}. \quad (96)$$

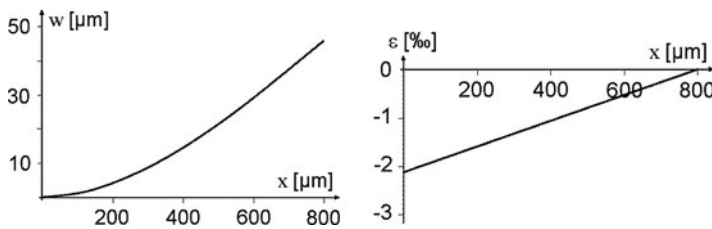
From this equation, the elastic force  $F_B$  of a beam deflected for  $w_0$  can be calculated also:

$$F_B = - \frac{3 E_B I}{L_B^3} w_0. \quad (97)$$

This equation is the characteristic curve of a beam clamped at one end and loaded with a force at the other end. It describes the elastic force generated by the beam as a function of its deflection.

**Table 3** Area moments of inertia of beams as a function of their cross-section. The position of the neutral fiber is  $d_{B,n}$

Form of cross-section		Area momentum of inertia
Rectangle		$\frac{1}{12} b_B d_B^3$
Circle		$\frac{1}{4} \pi R_B^4$
Trapezoid		$\frac{d_B^3}{36} \frac{b_{B,1}^2 + 4 b_{B,1} b_{B,2} + b_{B,2}^2}{b_{B,1} + b_{B,2}}$ $d_{B,n} = \frac{d_B}{3} \frac{2 b_{B,1} + b_{B,2}}{b_{B,1} + b_{B,2}}$
Unisotropically etched silicon		See above with $b_{B,1} = b_{B,2} + 1.416 d_B$



**Fig. 46** Deflection (*left*) and strain (*right*) of a microbeam as a function of the position along the beam

The strain  $\varepsilon_B$  on the surface of a beam clamped at one end is derived from (94) with (5) on page 7:

$$\varepsilon_B = \frac{d_B}{2} \frac{F (L_B - x)}{E_B I}. \tag{98}$$

Figure 46 shows the deflection and the strain on the surface of an 800- $\mu\text{m}$  long, 40- $\mu\text{m}$  wide, and 20- $\mu\text{m}$  thick beam with a Young’s modulus of 140 GPa clamped at one end, which is loaded transversally with a force of 1 mN at its free end.

Equation (98) and Fig. 46 show that the absolute value of the strain is maximal next to the clamped end of the beam. In a beam clamped at one end only, there is no residual stress; therefore, stress does not affect the bending of such a beam. This is good for building sensors, because stress is a parameter which is hard to control. Stress often changes with temperature or as a function of the packaging or fixation of a sensor. Thus, it is important to note that a beam clamped at only one end is a basic element of microtechnique, which is independent of stress.

Temperature changes still have a small effect on the deflection of a beam fixed at one end, because Young's modulus is a weak function of temperature.

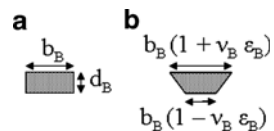
Equations (96) and (98) show also that deflection and strain of a beam clamped at one end are linear functions of the force acting at the free end of the beam. This is also useful for building sensors, because a linear characteristic curve is desirable in many cases.

If a beam shall be used as a sensor, it is necessary to measure its deflection. A possible way to do this is to employ *strain gauges* on a surface of the beam which is strained during deflection. The strain in the direction of the beam is described by (98), and the strain perpendicular to this can be calculated from Poisson's ratio  $\nu_B$  of the beam. Figure 47 shows exaggeratedly how the cross-section of a beam is changing when it is bent upward. The beam surface that is stretched by  $\varepsilon_B$  narrows in the two perpendicular directions (width and thickness) by  $\nu_B \varepsilon_B$  and the opposite surface which is compressed longitudinal widens transversally by the same amount.

As in the case of the membranes, it is assumed now that the strain gauges are so thin that they do not hinder the bending and straining of the beam. This assumption is not correct in most cases, but it eases the calculation a lot and the mechanical effect of the strain gauges just diminishes the resistance change measured in reality and does not cause principally different results. When metal strain gauges on the beam are used (geometrical effect only; cf. page 21f), the resistance change of strain gauges longitudinal  $\Delta R_{el,l}$  and transversal  $\Delta R_{el,t}$  to the beam can be calculated now with (78) on page 55. It is simply necessary to replace the strain of the membrane by the strain of the beam:

$$\begin{aligned} \frac{\Delta R_{el,l}}{R_{el,l}} &= \varepsilon_l(1 + \nu) - \varepsilon_t(1 - \nu) = \varepsilon_B(1 + \nu) + \nu_B \varepsilon_B(1 - \nu) \\ &= \varepsilon_B(1 + \nu_B + \nu(1 - \nu_B)), \end{aligned} \quad (99)$$

$$\begin{aligned} \frac{\Delta R_{el,t}}{R_{el,t}} &= \varepsilon_l(1 + \nu) - \varepsilon_t(1 - \nu) = -\nu_B \varepsilon_B(1 + \nu) - \varepsilon_B(1 - \nu) \\ &= -\varepsilon_B(1 + \nu_B - \nu(1 - \nu_B)). \end{aligned} \quad (100)$$



**Fig. 47** Cross-section of a rectangular beam without (a) and with (b) bending upward

These equations show that the absolute value of the resistance change of a strain gauge longitudinal to the beam is a bit larger than for a transversal one. The sign of these resistance changes is different, and, therefore, they can be combined to a Wheatstone bridge similar as in the case of strain gauges on a membrane described on page 53. Figure 48 shows for the beam from Fig. 46 the resistance changes of strain gauges, as calculated with (99) and (100). The strain  $\epsilon_B$  of the beam as a function of the position  $x$  in longitudinal direction was calculated with (98) and the area momentum of inertia of a beam with rectangular cross-section from Table 3. This figure shows that the typical resistance change which can be expected is on the order of some  $10^{-3}$  and, therefore, temperature compensation by a bridge is required. The sensitivity is larger, if the space next to the clamped end of the beam is used for the longitudinal strain gauges and the transversal ones are arranged adjacent to them, because longitudinal strain gauges show the larger resistance change.

It is important to mention here that the resistance change of strain gauges doped into single crystalline silicon are a function of crystal orientation and need to be calculated with (21) (on page 22) instead of (99) and (100).

The transversal deflection of a *beam loaded with its own weight* is calculated with the following equation:

$$w = \frac{\rho_B g_e b_B d_B}{24 E_B I} x^2 (x^2 - 4 L_B x - 6 L_B^2). \tag{101}$$

In this equation,  $g_e$  and  $\rho_B$  are the gravitational acceleration of the earth ( $g_e \approx 9.81 \text{ m/s}^2$ ) and the density of the beam, respectively. This equation shows that the deflection of a beam loaded with its own weight is proportional to the square of its linear dimensions. Thus, if length, thickness, and width are reduced by a factor 10, the deflection is decreased by a factor 100. The Golden Gate Bridge at San Francisco is 2.7 km long and approximately 10 m thick, Young’s modulus, Poisson’s ratio, and the density of steel are 210 GPa, 0.28, and  $7,900 \text{ kg/m}^3$ , respectively. It can hold its own weight only because it is supported by ropes being fixed at the two towers (cf. Fig. 49). If all linear dimensions of the bridge would be reduced to a millionth, i.e., 2.7 mm long and  $10 \text{ }\mu\text{m}$  thick, according to (101) it would hang down by 930 nm only. That is, in the microworld, the Golden Gate Bridge can be built as a

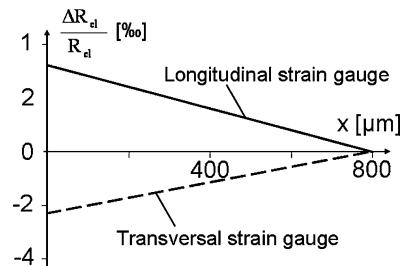
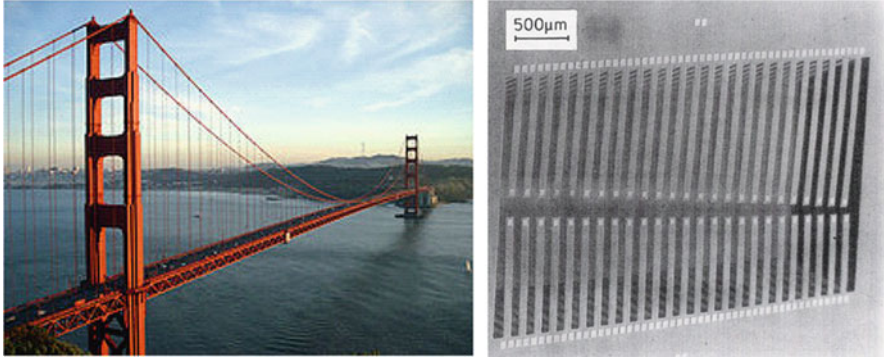
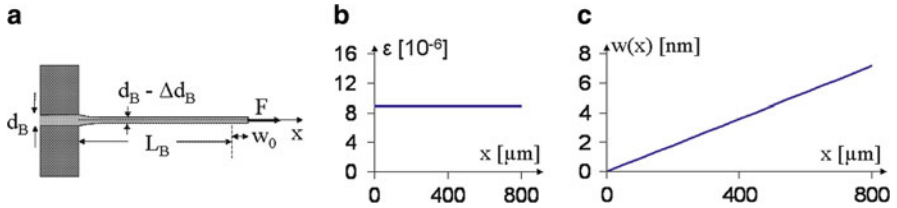


Fig. 48 Resistance change of strain gauges on a beam



**Fig. 49** The deflection of beams loaded with their own weight is much smaller in the microscopic than in the macroscopic world. The Golden Gate Bridge (on the *left*) [35] needs to be supported by ropes, while extreme cantilevers [44] (on the *right*) can be build without such supports in microtechnique



**Fig. 50** Beam deflected in longitudinal direction (a), strain (b), and deflection (c)

simple beam without any extra-supporting structure. Figure 49 shows an extreme cantilever structure which is possible in the microscopic world only.

*Beams can be deflected in longitudinal direction* also (cf. Fig. 50a). If a force is acting in longitudinal direction, a beam is deflected according to Hooke’s law:

$$\epsilon_B = \frac{\sigma_B}{E_B} = \frac{F}{A_B E_B}. \tag{102}$$

In this equation,  $\sigma_B$ ,  $E_B$ ,  $A_B$ , and  $F$  are the mechanical stress in the beam, Young’s modulus, the cross-sectional area of the beam, and the force acting at the end of the beam, respectively. The deflection  $w$  of the beam is the product of the distance from the clamping point  $x$  and the strain  $\epsilon_B$ :

$$w(x) = \epsilon_B x = \frac{F}{A_B E_B} x. \tag{103}$$

Figure 50b, c shows strain and deflection of the beam assumed for Fig. 46 but loaded with an axial force of 1 mN at the free end.

A longitudinal force acting on a beam besides the longitudinal strain will cause transversal strains  $\varepsilon_t$  which are a function of Poisson's ratio of the beam:

$$\varepsilon_t = -\nu_B \varepsilon_B = -\nu_B \frac{F}{A_B E_B}. \quad (104)$$

As in the case of a beam clamped at one end and loaded transversally at the free end, the resistance change of metal strain gauges on the surface of the beam can be calculated now with (16) on page 21:

$$\Delta R_{el,l} = R_{el}(\varepsilon_L - \varepsilon_b - \varepsilon_d) = R_{el} \varepsilon_B(1 + \nu_B + \nu), \quad (105)$$

$$\Delta R_{el,t} = R_{el}(\varepsilon_L - \varepsilon_b - \varepsilon_d) = R_{el} \varepsilon_B (-\nu_B - 1 + \nu) = -R_{el} \varepsilon_B(1 + \nu_B - \nu). \quad (106)$$

Equations (105) and (106) are identical with (99) and (100), which describe the resistance change of a beam loaded transversally. However, the strain on the surface of a beam loaded transversally is much larger than the one of a longitudinally loaded beam (cf. Figs. 46 and 50, respectively). Therefore, if the beam is an elastic element in a sensor, e.g., an acceleration sensor, it should be designed such that it is loaded transversally. If the beam is employed to hold an object in position, however, a longitudinal load is preferable.

Note that the resistance change is to be calculated with the equations shown here only for metal strain gauges. Often strain gauges from silicon are used and their resistance change needs to be calculated as a function of their crystal orientation as shown in Fig. 17 on page 25. It needs to be considered also that Young's modulus and Poisson's ratio are a function of crystal orientation, if monocrystalline silicon is used to construct a beam (cf. Fig. 38 on page 50). Since it is difficult to take this into account in an analytical calculation, mean values are used often, which provide an approximate description.

If a beam clamped or fixed at one end is compressed by a force acting at the free end, the beam tends to bend transversally when a certain compressive stress is overcome. This force is called the *critical force*  $F_K$ . It is a function of the boundary condition at the ends of the beam and can be found in text books on mechanics [32–34]. Table 4 shows the cases which are most important in microtechnique.

A *beam clamped at both ends* and loaded with a transversal force  $F$  in its center (cf. Fig. 51a) is another basic element which is important in microtechnique. The deflection  $w$  of this beam as a function of the distance  $x$  from the beam center, its length  $L_B$ , width  $b_B$ , thickness  $d_B$  (dimension in the direction of the deflection), and Young's modulus  $E_B$  is described by the following equation [33]:





$$w = \frac{F}{48 E_B I} \left( \frac{1}{4} L_B^3 - 3 L_B x^2 + 4|x^3| \right). \quad (107)$$

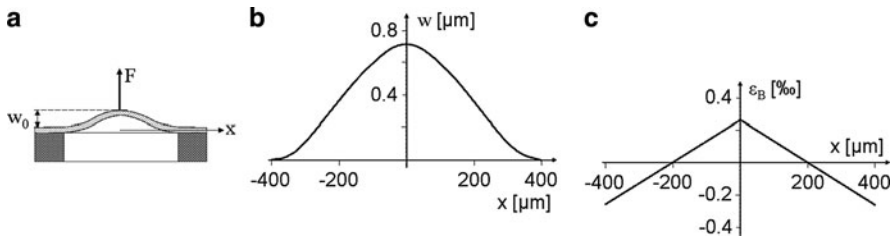
The maximum deflection  $w_0$  occurs at the center ( $x = 0$ ):

$$w_0 = \frac{L_B^3 F}{192 E_B I}. \quad (108)$$



**Table 4** Critical force (buckling load) of beams with different boundary conditions

	$F_K = -\frac{\pi^2 E_B I}{4 L_B^2}$
	$F_K = -\pi^2 \frac{E_B I}{L_B^2}$
	$F_K = -2 \pi^2 \frac{E_B I}{L_B^2}$
	$F_K = -4 \pi^2 \frac{E_B I}{L_B^2}$



**Fig. 51** Beam clamped at both ends and loaded transversally at its center. (a) Schematic drawing, (b) deflection, and (c) strain

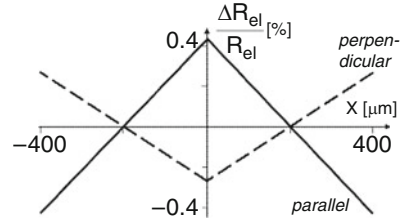
The strain  $\varepsilon_B$  due to bending of such a beam is on its surface according to (5) on page 7 and (107):

$$\varepsilon_B = \frac{d_B F}{16 E_B I} (L_B - 4 |x|). \tag{109}$$

Deflection and strain calculated with (107) and (109), respectively, of a beam clamped at both ends and loaded with a force of 1 mN at its center which has the same properties as the beam in Fig. 46 (page 67) are shown in Fig. 51. The comparison with Fig. 46 and of (108) with (96) (on page 66) shows that the deflection of a beam clamped at both ends is 64 times smaller than the same beam clamped at one end only. This shows that a beam clamped only at one end is superior as an elastic element in a sensor, but is less suitable as a supporting element.

The resistance change of metal strain gauges on the surface of a beam clamped at both ends is calculated with (99) and (100) again, because they undergo the same change as a function of the strain of the beam. However, (109) needs to be inserted for the strain at the surface of the beam. Figure 51c shows the strain distribution

**Fig. 52** Resistance change of metal strain gauges on a beam clamped at both ends parallel and perpendicular to the beam, respectively



along the surface of the beam calculated with this equation and Fig. 52 the resistance changes of strain gauges parallel and perpendicular to the beam, respectively.

Strain gauges are arranged best parallel to the beam, because the resistance change is larger this way. Opposite sign of the resistance change is achieved by placing them at the middle and at the end of the beam, respectively.

The equations above that describe the deflection and strain of beams were derived by considering the bending moments only. This is correct, if the beam is fixed or clamped only at one end, because no stress along the neutral fiber of the beam can develop; and this is sufficient in the macroscopic world, where, e.g., the beams employed to erect a building are designed such that their deflection is much less than their thickness. In microtechnique, however, beam deflections occur which are much larger than their thickness, and, if these beams are clamped at both ends, the strain and stress of the neutral fiber need to be taken into account.

This is similar as in the case of thin membranes which was discussed above (cf. page 32). Therefore, the stress  $\sigma_D$  due to straining of the neutral fiber is calculated in a similar way from the increase in length which is necessary to deflect the beam [cf. (31), page 32]:

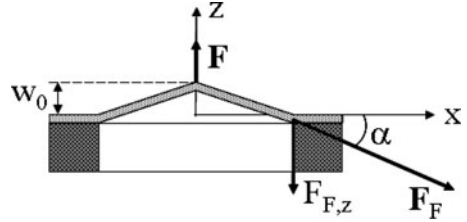
$$\sigma_D = \frac{8}{3} \frac{w_0^2}{L_B^2} E_B. \quad (110)$$

The total stress along the neutral fiber is the sum of the stress due to straining of the neutral fiber and the residual stress which is nearly always present in microtechnique:

$$\sigma = \sigma_0 + \sigma_D = \sigma_0 + \frac{8}{3} \frac{w_0^2}{L_B^2} E_B. \quad (111)$$

The equilibrium of forces at the ends of the beam is shown in Fig. 53. The beam is deflecting until the force of the frame  $F_F$  is compensating the force  $F$  pulling the beam upward. The x-components of the two forces acting at the frame cancel out each other. As a consequence, it is sufficient to calculate the z-components of these forces  $F_{F,z}$  which must be equal to the absolute value of the force  $F$  [cf. (112)]. The

**Fig. 53** Equilibrium of forces at the ends of a beam



force of the beam which is equal to the forces of the frame  $F_F$  can be expressed as the tensile stress  $\sigma$  of the beam times its cross-section  $A_B$ . The  $z$ -component of this force is obtained by multiplying with  $\sin(\alpha)$ . For small deflections, the sine is equal to the slope which is the ratio of  $z$ - and  $x$ -components.

$$F = 2 F_{F,z} = 2 \sigma A_B \sin(\alpha) \approx 2 \sigma A_B \frac{w_0}{L_B/2} = 4 \frac{A_B}{L_B} w_0 \sigma. \quad (112)$$

The stress of the beam is now inserted in (112) from (111) and an equation for the interrelationship of the deflection of a beam clamped at both ends and the force acting at its center is obtained:

$$F = 4 \frac{A_B}{L_B} w_0 \left( \sigma_0 + \frac{8}{3} E_B \frac{w_0^2}{L_B^2} \right). \quad (113)$$

This equation approximately describes the effect of the stress of the beam on its deflection, but the effect of the bending moments is not taken into account. As a rough approximation, the forces due to stress and bending moments (108) are added now, resulting in an equation which includes all important effects:

$$F = 4 \frac{A_B}{L_B} w_0 \left( 48 \frac{E_B I}{A_B L_B^2} + \sigma_0 + \frac{8}{3} E_B \frac{w_0^2}{L_B^2} \right). \quad (114)$$

This equation is only a rough approximation. If exact calculations are required FEM will help to find the desired results. On the other hand, (114) contributes to the understanding of beam deflections very much. The three terms in the parentheses correspond to the contributions of bending moments, residual stress, and stress due to straining, respectively. Similar as for a membrane (cf. page 41), there exists a critical compressive stress  $\sigma_k$  at which the beam buckles up, and this critical stress is reached when the residual stress in (114) is equal to the negative of the term of the bending moments:

$$\sigma_k = -48 \frac{E_B I}{A_B L_B^2}. \quad (115)$$

When the residual stress is more compressive than the critical stress, the beam is *bistable* and can bulge either up or down such as a membrane with compressive stress (cf. page 39). A bistable beam can retain a bulging direction until the counter force overcomes a certain level and the beam flips over to the other side. Figure 54 shows the elastic force calculated with (114) for a rectangular beam with the same properties as used above in Figs. 46 (on page 67) and 51 but clamped at both ends and with a residual stress of 0 MPa and -600 MPa, respectively.

The deflection of such a beam not loaded by a force is calculated from the zeros of (114). One solution is found for zero deflection which corresponds to the origin of the graph in Fig. 54. Two other solutions are found by the zero of the term in the parentheses of (114):

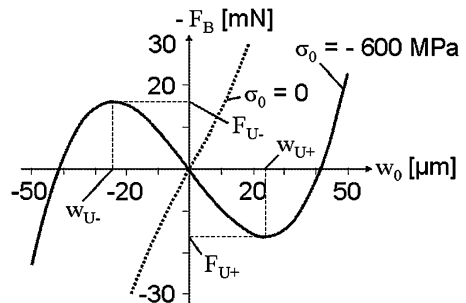
$$\begin{aligned} w_0(F=0) &= \pm 6 \sqrt{\frac{2I}{A_B}} \sqrt{\frac{\sigma_0}{-48 (E_B I / (A_B L_B^2))} - 1} \\ &= \pm 3\sqrt{2} \sqrt{\frac{I}{A_B}} \sqrt{\frac{\sigma_0}{\sigma_k} - 1}, \end{aligned} \quad (116)$$

where  $\sigma_k$  is the critical stress of the beam as calculated with (115). The deflection  $w_0$  of a beam clamped at both ends without any load as a function of its residual stress (116) is similar as the deflection of a membrane without any pressure drop as a function of its residual stress [(57) on page 42]), and, therefore, the graph of this interrelationship is similar to that of Fig. 29 on page 42.

The deflection  $w_U$  at which the beam snaps over to the opposite side is calculated from the zeros of the derivatives of (114):

$$\frac{\partial F}{\partial w_0} = 0 = \frac{4 A_B}{L_B} \left( 48 \frac{E_B I}{L_B^2} + \sigma_0 + 8 E_B \frac{w_0^2}{L_B^2} \right) \Rightarrow w_U = \pm \sqrt{6} \sqrt{\frac{I}{A_B}} \sqrt{\frac{\sigma_0}{\sigma_k} - 1}, \quad (117)$$

where  $\sigma_k$  is the critical stress of the beam again. The comparison of (116) and (117) shows that the ratio of the deflection without load to the deflection at the snapping over is  $1/\sqrt{3} \approx 0.58$  independent of the other parameters of the beam. This is the same as for membranes (cf. page 43).



**Fig. 54** Elastic force of a beam as a function of its deflection and residual stress as calculated with (114)

If there is no force acting on the beam and it is bending downward, its position can be calculated with (116). In the graph shown in Fig. 54, this position is approximately  $-42 \mu\text{m}$ . If the beam is to be moved towards its non-deflected position, the force shown in the graph needs to be applied. This force increases until  $w_U$  is reached. If the beam is further moved upward, the force needed to hold it in that position decreases, while the non-deflected position at the origin in Fig. 54 is approached. No force is necessary to hold the beam in the non-deflected position, but if it is moved away from that position, a certain force is required to prevent it from moving further away. This is a typical unstable position.

The force required to let the beam snap over to the opposite side is calculated by inserting the position of snapping over [(117) with the critical stress from (115)] into (114) which describes the force of the beam as a function of its deflection:

$$F_U = \pm 8 \sqrt{\frac{2}{3} \frac{\sqrt{A_B I}}{L_B}} \sigma_k \left( \frac{\sigma_0}{\sigma_k} - 1 \right)^{3/2}. \quad (118)$$

The exact values of the critical force  $F_K$  at which a beam buckles was shown in Table 4 (page 72). The critical stress can be calculated from this by dividing the cross-section of the beam. This results in:

$$\sigma_k = -4\pi^2 \frac{E_B I}{A_B L_B^2} = -39.5 \frac{E_B I}{A_B L_B^2}. \quad (119)$$

This is approximately 20% less than the value calculated with (115) and provides an impression of the effect of the rough approximations made when (114) was derived.

Two beams may be arranged such that they are *deflected in parallel* (cf. Fig. 55). This results in a parallel movement of a structure fixed at these beams. The comparison of Figs. 55 and 51a (page 72) shows that the deflection curve of each of the parallel beams is half of the deflection curve of a beam clamped at both ends. Thus, the deflection of both arrangements should be the same with the only difference that the parallel beams have no residual stress and no stress due to straining of the neutral fiber. Besides this, it needs to be taken into account that the length of the beams in Fig. 55 is half of the beam length in Fig. 51a. Therefore,  $L_B$  needs to be replaced by  $2L_B$ , and the different origin of the coordinates needs to

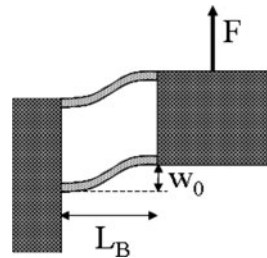


Fig. 55 Beams deflected in parallel

be taken into account in (107–109) to describe deflection  $w$  and strain  $\varepsilon_B$  at the surface as a function of the position along the beam, and maximum deflection  $w_0$ :

$$w = \frac{F}{24 E_B I} (3 L_B x^2 - 2 x^3), \quad (120)$$

$$w_0 = \frac{L_B^3 F}{24 E_B I}, \quad (121)$$

$$\varepsilon_B = \frac{d_B F}{8 E_B I} (L_B - 2 x). \quad (122)$$

Deflection and strain due to bending are the same as in the left half of the graphs in Fig. 51b, c (page 72).

Another combination of beams employed very often is two beams arranged in parallel with their ends fixed to each other (cf. Fig. 56a). One of the beams is supported at its center and the load  $F_t$  is applied at the center of the other beam in transversal direction (Fig. 56b). This arrangement may be regarded as a *folded beam*. It simply doubles the displacement of a single beam clamped at both ends and is not subject to residual stress and stress due to straining of the neutral fiber [(108) on page 71].

Folded beams allow a larger displacement with a given force and provide a strait movement similar to beams deflected in parallel. They are very stiff when they are pulled by a longitudinal force  $F_l$ , i.e., perpendicular to the intended direction (Fig. 56c). Therefore, folded beams are very popular for the movable electrodes of electrostatic comb actuators which are not able to produce large forces and are sensitive to displacements in transversal direction (cf. page 132). The displacement  $w_1$  in longitudinal direction of a folded beam with length  $L_B$ , Young's modulus  $E_B$ , area moment of inertia  $I$ , and a length  $L_{Bc}$  of the beams connecting the two parts of the folded beam can be estimated by the following equation, which was derived from finite element calculations (FEM) [31]:

$$F_l = (49 \pm 5) \frac{E_B I}{L_{Bc}^2 L_B} w_1. \quad (123)$$

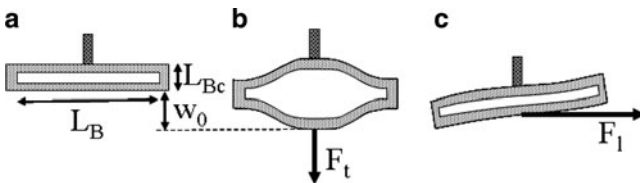
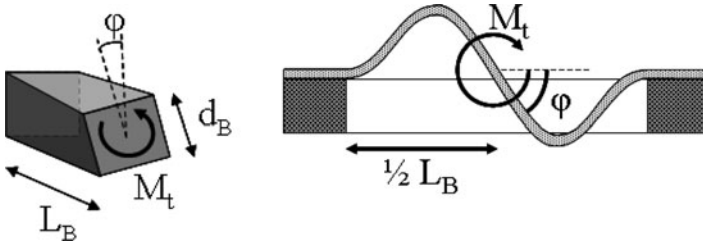


Fig. 56 Folded beams with (b) and without (a) deflection



**Fig. 57** Beams twisted by axial (*left*) and transversal (*right*) torques

**Table 5** Torsional constants of beams

Circular cross-section	$I_t = \frac{\pi}{2} R_B^4$				
Square cross-section	$I_t = 0.1406 d_B^4$				
Rectangular cross-section	$I_t = a \frac{d_B^3 b_B^3}{d_B^2 + b_B^2}$				
$d_B/b_B$	1	2	4	8	$\infty$
$a$	0.281	0.286	0.299	0.312	1/3

The standard deviation expresses how good the FEM results are described by the above equation. The FEM calculations also showed that the deflection in longitudinal direction is not a function of the deflection in transversal direction.

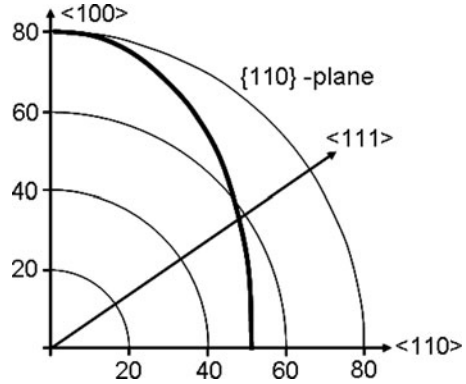
A twisted beam shows *torsional deflection* (cf. Fig. 57 left). Since torsional elastic forces are comparatively small for microbeams, torsional deflection is important in applications where only small forces are available and needed, e.g., sensors or the steering of light beams. The angle  $\varphi$  at which a beam with the length  $L_B$  is twisted by a torque  $M_t$  is a function of its torsional constant  $I_t$  and the shear modulus  $G_B$  of its material:

$$\varphi = \frac{L_B}{G_B I_t} M_t. \quad (124)$$

The torsional constant is found in text books, e.g., [32–34] and is shown in Table 5 for beams with three different cross-sectional shapes. The shear modulus is a material constant and found in text books also or is calculated from Young's modulus and Poisson's ratio:

$$G_B = \frac{E_B}{2(1 + \nu_B)}. \quad (125)$$

**Fig. 58** Shear modulus [GPa] of monocrystalline silicon. Reprinted with permission from [29]. Copyright [1965], American Institute of Physics



All the equations above are based on the assumption that Young's modulus and Poisson's ratio are isotropic quantities. As shown in Fig. 38 on page 50, this is not true for monocrystalline materials such as silicon. As a consequence, the shear modulus is a function of crystal orientation also. This interrelationship is shown in Fig. 58.

A beam clamped at both ends can be twisted perpendicular to its axis also (cf. Fig. 57 right). Normally, this is not desired, but the designer of a microcomponent needs to know also, which not intended deflections may occur at basic elements.

The deflection angle of a beam with an area momentum of inertia  $I$ , clamped at both ends, and subject to a torque  $M_t$  perpendicular to its axis at its center is:

$$\varphi = \frac{L_B}{16 E_B I} M_t. \quad (126)$$






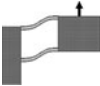

The effect of residual stress is not included in this equation and further reduces the deflection angle.

Table 6 summarizes the interrelationships of deflections and forces and torques. The comparison of the equations in this table shows that comparatively large forces can be achieved with membranes and with beams clamped at both ends. Thus, these elements are preferable for stable supports and actuators. Large deflections are easily achieved with beams clamped at one end only. Therefore, these are superior for elastic elements in sensors and movable supports and bearings in microtechnique.

All beams clamped at one end only are not affected by stress and stress changes. Stress changes most often are generated by temperature changes because materials with different thermal extensions are involved. Thus, these beams are less sensitive to temperature changes which are the source of most cross sensitivities. This is


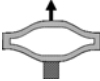



**Table 6** Equations for the calculation of the interrelationship of deflection  $w_0$  and force  $F$  or torque  $M_t$  of a beam or membrane

Type	Equation
General solution for circular membrane, rough approximation	$F = 4 \pi d_M w_0 \left( \frac{4}{3} \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64}{105} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right)$
Beam, general solution, rough approximation	 $F = 4 \frac{A_B}{L_B} w_0 \left( 48 \frac{E_B I}{A_B L_B^2} + \sigma_0 + \frac{8}{3} E_B \frac{w_0^2}{L_B^2} \right)$
Beam, without stress, exact solution	 $F = 192 \frac{E_B I}{L_B^2} w_0 \Rightarrow w_0 = \frac{L_B^2}{192 E_B I} F$
Exact	 $F = 3 \frac{E_B I}{L_B^3} w_0 \Rightarrow w_0 = \frac{L_B^3}{3 E_B I} F$
Under own weight, exact	 $w_0 = - \frac{3}{8} \frac{\rho_B g e b_B d_B}{E_B I} L_B^2$
Exact	 $F = \frac{A_B E_B}{L_B} w_0 \Rightarrow w_0 = \frac{L_B}{A_B E_B} F$
Exact	 $F = 24 \frac{E_B I}{L_B^2} w_0 \Rightarrow w_0 = \frac{L_B^3}{24 E_B I} F$
Exact	 $M_t = \frac{G_B I_t}{L_B} \phi \Rightarrow \phi = \frac{L_B}{G_B I_t} M_t$

(continued)

**Table 6** (continued)

Type	Equation
Beam without stress	
 Exact	$M_t = \frac{16 E_B I_t}{L_B} \phi \Rightarrow \phi = \frac{L_B}{16 E_B I_t} M_t$
 Approximation	$F_t = 96 \frac{E_B I}{L_B^3} w_0 \Rightarrow w_0 = \frac{L_B^3}{96 E_B I} F_t$
 Approximation	$F_t = (49 \pm 5) \frac{E_B I}{L_{Bc}^2 L_B} w_1 \Rightarrow w_1 = \frac{L_{Bc}^2 L_B}{(49 \pm 5) E_B I} F_t$

For beams from monocrystalline material, the Young's modulus in longitudinal direction has to be used (cf. Fig. 38 on page 50). For membranes from monocrystalline materials using a mean value of Young's modulus yields an approximate result

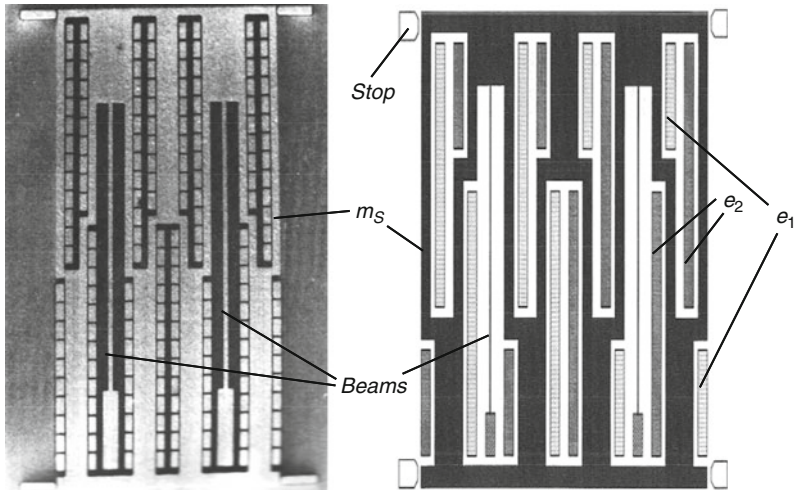
another reason why beams clamped at only one end are preferable in a lot of applications.

## Exercises

### Problem 12

Figure E13 shows an acceleration sensor fabricated from nickel with the LIGA process [36]. For the measurement, the capacity between the fixed electrodes  $e_1$  and  $e_2$  and the seismic mass  $m_s$  is recorded. As shown in the schematic drawing of the figure, the seismic mass is suspended by parallel beams to achieve a homogeneous width of the capacitor gap when the beams are deflected. At the sides, there are stops mounted which prevent a short circuit when an unexpectedly large acceleration occurs. The seismic mass may move up to  $3.5 \mu\text{m}$  until it touches one of the stops.

- What is the maximum acceleration which can be measured with the sensor?
- What is the maximum strain which can occur in the beams?
- What is the minimum height at which the beams need to be mounted above the substrate not to be deflected down by their own weight and touching the substrate?



**Fig. E13** Acceleration sensor fabricated by the LIGA process shown as an SEM (*left*) and as a schematic drawing (*right*) [36]. (Courtesy of Karlsruhe Institute of Technology, KIT)

Hint: Assume that the entire seismic mass is concentrated at the end of the beams.

Young's modulus of nickel	186 GPa	Height of seismic mass and beams	80 $\mu\text{m}$
Length of the beams	2.5 mm	Width of the beams	14.5 $\mu\text{m}$
Seismic mass	1.3 mg		

### **Problem 13**

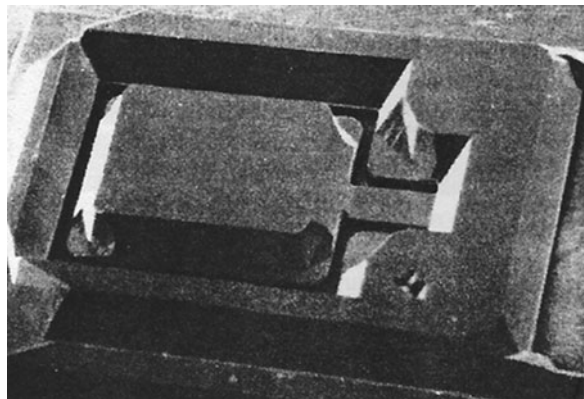
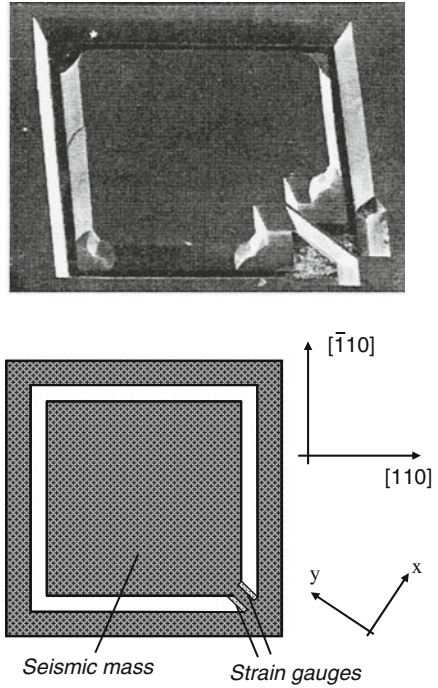
Figure E14 shows at the top a silicon acceleration sensor which was etched from a (100)-wafer and consists of a seismic mass which is suspended by a beam with a nearly rectangular cross-section [37]. On the upper surface of the beam, two strain gauges are integrated which are combined to a Wheatstone bridge with two reference resistances (cf. Fig. E14 bottom).

The strain gauges of the beam are fabricated from n-doped conductor paths and are parallel to the  $\langle 100 \rangle$ -direction. The beam and the seismic mass consist of p-doped silicon.

The beam is 600  $\mu\text{m}$  long and 40  $\mu\text{m}$  wide. The width of the strain gauges is 10  $\mu\text{m}$  and they stretch on the beam next to its edge. They are as long as the beam. The length of the edge of the seismic mass is 3 mm and both the seismic mass and the beam are 250  $\mu\text{m}$  high.

- (a) How does this sensor work? What could be the advantages compared with the usual designs of acceleration sensors shown in Fig. E15?

**Fig. E14** SEM and schematic drawing of the acceleration sensor [37].  
 © [1997] IEEE



**Fig. E15** The first silicon acceleration sensor [38].  
 © [1979] IEEE

- (b) The sensor is accelerated by 3 g in z-direction. What resistance change  $\Delta R_{ei}/R_{ei}$  develops in the strain gauges?  
 What output signal can be measured at the Wheatstone bridge?
- (c) Now the sensor is accelerated in x-direction and an output voltage of 0.3 mV is measured at the bridge. Please, calculate the acceleration. What resistance change could be measured at a single strain gauge?

Hints for the calculations:

- As an approximation, assume that the seismic mass would be concentrated in a point at the end of the beam.
- Besides this, you may assume that the seismic mass has a rectangular shape with the dimensions given in the table.

Height of the seismic mass	250 $\mu\text{m}$	Length of the edge of the seismic mass	3 mm
Density of silicon	2.32 $\text{g/cm}^3$	Young's modulus of silicon	160 GPa
Poisson's ratio of silicon	0.23	Specific resistance of the n-doped silicon	5 $\Omega\text{ cm}$
Voltage supply	5 V		

# Vibrations

Vibrations of membranes and beams are important in microtechnique. The frequency range of pressure sensors and microphones is limited by the resonance frequencies of their membrane. In a similar way, the resonance frequency of beams limits the possible applications of acceleration sensors. On the other hand, the resonance frequency of a vibrating element may be proportional to the measurand and allow measurements less affected by noise.

Figure 59 shows the amplitude of a vibrating structure as a function of the exciting frequency. There are two frequency ranges which usually are important for the designer: The range well below the lowest resonance frequency – the so-called fundamental frequency  $f_1$ , and the range in the near of the fundamental frequency. The former range is desirable when a sensor or an actuator shall be independent of the frequency. For example, the deflection of the membrane of a pressure sensor shall not be enhanced by resonance effects. Therefore, pressure sensors are limited to a certain frequency range. On the other hand, actuators may achieve larger deflections when driven at their resonance and the resonance frequency may be employed to measure certain quantities such as forces and masses.

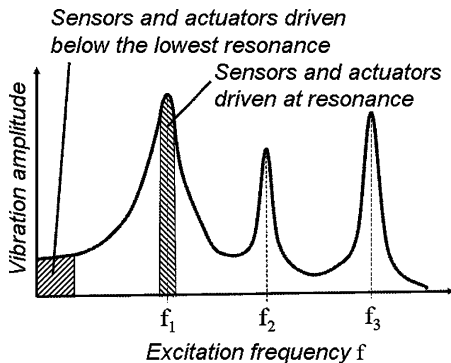
The classical way to calculate the resonance frequencies of a body is to set up the equilibrium of forces in the form of a differential equation and to solve that equation. For example, the differential equation for calculating the *vibrations of a membrane* neglecting bending moments and stress due to straining of the neutral fiber is:

$$\sigma_0 \left( \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right) = \rho_M \frac{\partial^2 w}{\partial t^2}. \quad (127)$$

The left side of this equation is the elastic force due to the residual stress  $\sigma_0$  of a rectangular membrane with a deflection  $w(x, y, t)$  which is a function of  $x$ - and  $y$ -coordinate and time  $t$ . On the right side of the equation is the inertial force according to Newton's law where the mass is represented by the density  $\rho_M$  of the membrane. This differential equation is solved by the following ansatz:

$$w_{n,m}(x, y, t) = A_{n,m} \sin\left(\frac{n\pi}{a_M} \left[x + \frac{a_M}{2}\right]\right) \sin\left(\frac{m\pi}{b_M} \left[y + \frac{b_M}{2}\right]\right) \sin(\omega_{n,m} t + \Delta\varphi_{n,m}). \quad (128)$$

**Fig. 59** Vibration amplitude as a function of excitation frequency



In this ansatz,  $a_M$  and  $b_M$  denote the length of the membrane in  $x$ - and  $y$ -direction, respectively, and  $A_{n,m}$ ,  $\omega_{n,m}$ , and  $\Delta\varphi_{n,m}$  are the amplitude, angular frequency, and phase of the vibration in the mode  $n,m$ , respectively. This ansatz can fulfill the above differential equation only if the frequency  $f$  adopts one of the following values:

$$f_{n,m} := \frac{\omega_{n,m}}{2\pi} = \frac{1}{2} \sqrt{\frac{\sigma_0}{\rho_M}} \sqrt{\frac{n^2}{a_M^2} + \frac{m^2}{b_M^2}}. \tag{129}$$

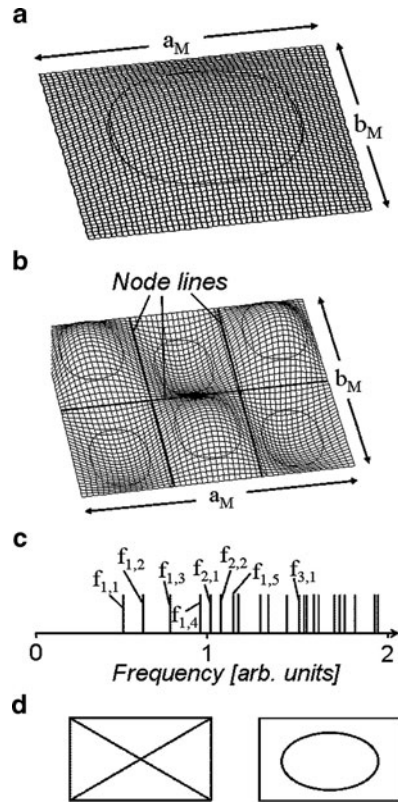
These values are the resonance frequencies of the membrane. The lowest possible frequency  $f_{1,1}$  with  $n = m = 1$  is the fundamental frequency. Each pair of natural numbers  $n$  and  $m$  denote a so-called vibration mode of the membrane and every mode corresponds to a certain resonance frequency. Figure 60 shows a deflected membrane vibrating at its fundamental frequency  $f_{1,1}$  and at  $f_{3,2}$ . The numbers  $n$  and  $m$  correspond to the numbers of antinodes in  $x$ - and  $y$ -direction, respectively. In Fig. 60c, there are shown the resonance frequencies calculated with (129) with  $a_M = 1$  and  $b_M = 2.4$ . There are even more resonances, because there are further solutions of (127) which are not included in the ansatz (128). The node lines (lines which are not deflected during vibration) of such resonances are shown in Fig. 60d.

In general, a membrane will vibrate at several resonance frequencies at the same time. The general solution of (127) is the sum of all solutions possible. If we restrict ourselves to solutions according to the ansatz (128), the general deflection  $w(x, y, t)$  is given by:

$$w(x, y, t) = \sum_{n,m} w_{n,m}(x, y, t). \tag{130}$$

The variety of frequencies that appear above the fundamental frequency impede employing higher resonances for sensors and actuators, because it is difficult to make sure that the desired resonance is in use and not a neighboring one. Therefore,

**Fig. 60** Vibrations of rectangular membranes: (a) fundamental frequency  $f_{1,1}$ ; (b)  $f_{3,2}$ ; (c) frequency spectrum calculated with (129); and (d) additional nodal lines



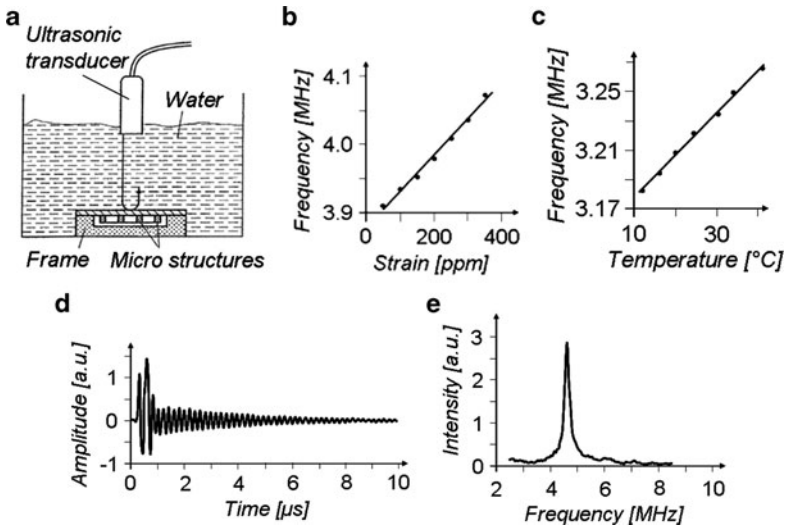
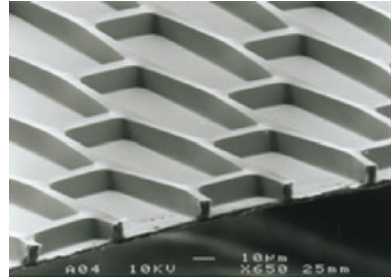
in most applications in microtechnique, it is enough to know the fundamental frequency and the effects which may change it.

Equation (129) shows that, besides the geometry, the fundamental frequency is a function of membrane stress and density, i.e., changes of the geometry, density, and stress can be determined by measuring the fundamental resonance frequency of a membrane. On the other hand, changes of these properties may interfere measurements of this kind. An example is *micromembrane sensors* [39]. Figure 61 shows a break through such a sensor. A membrane is separated by a microstructure into micromembranes which due to their small dimensions show a fundamental resonance frequency of some MHz, which is in the range of ultrasound used for medical diagnosis.

The micromembranes can be excited to vibrations by an acoustic pulse generated by an ultrasonic transducer. After this, the membranes emit ultrasound at their resonance frequency. Figure 62a shows the setup of an experiment in which micromembranes were placed on the bottom of a vessel filled with water. The micromembranes were excited by a transducer pulse and the same transducer picked up the acoustic response from the membranes which is shown in Fig. 62d.



**Fig. 61** Break through a micromembrane sensor. (Courtesy of Karlsruhe Institute of Technology, KIT)



**Fig. 62** Experimental test setup for micromembrane sensors [39] (a); Measured resonance frequency as a function of strain (b) and temperature (c); Measured signal from the micromembranes (d); and Fourier transformation of the second part of that signal (e). (Courtesy of Karlsruhe Institute of Technology, KIT)

The first part of the signal is the excitation pulse reflected from the membranes and the second part of the signal is the sound emission from the micromembranes. From this second part of the signal, a Fourier transformation was calculated which yields the resonance frequency (cf. Fig. 62e).

According to (129), the resonance frequency is a function of membrane stress which may be a function of temperature also, if the thermal expansion of the housing of the membranes is different from the one of the membranes. Figure 62b, c shows the resonance frequency measured as a function of straining and temperature [39].

When (129) was derived, residual stress was taken into account only. Stress due to straining of the neutral fiber and bending moments influences the resonance frequency also. Therefore, this equation is only correct for thin membranes with comparatively small deflections and large residual stress. The stress due to straining

of the neutral fiber is only important for membranes vibrating with large amplitudes. This does not occur often, and, therefore, here is not discussed further. The bending moments are included employing the *Rayleigh method* [40]. Similar to the Ritz method (cf. page 34), the energy of the vibrating membrane is calculated. When an elastic body or structure is vibrating, its energy is transformed back and forth between potential and kinetic energy. If damping is neglected, the maxima of potential and kinetic energy are equal. From this equation, the resonance frequency can be calculated.

Figure 63 shows a simple example. A mass affixed to the end of a spring (spring constant  $k$ ) is vibrating around its rest position. The maximum potential energy  $V_{p,m}$  of this system is reached at the extreme positions of the deflections where the velocity is zero and the maximum kinetic energy  $E_{kin,m}$  is achieved while passing its idle position with maximum velocity  $v_{max}$ :

$$V_{p,m} = \frac{k}{2} w_0^2 = E_{kin,m} = \frac{m_K}{2} v_{max}^2. \tag{131}$$

The position of every vibrating body is described by an equation of the following type:

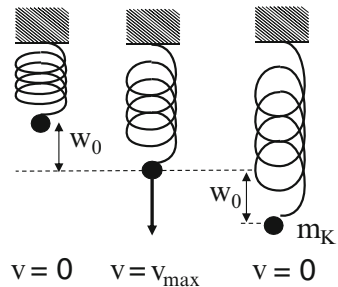
$$w(t) = w_0 \sin(\omega t + \varphi) \Rightarrow v(t) = \frac{\partial w}{\partial t} = w_0 \omega \cos(\omega t + \varphi). \tag{132}$$

The maximum of the velocity is reached when the cosine function equals 1. Therefore, (131) becomes now:

$$\frac{k}{2} w_0^2 = \frac{m_K}{2} w_0^2 \omega^2 \Rightarrow \omega = \sqrt{\frac{k}{m_K}}. \tag{133}$$

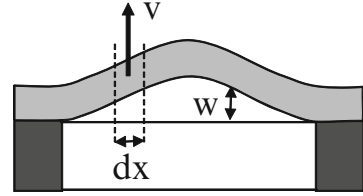
This result could have been obtained from the calculation of the equilibrium of forces with the inertial force of Newton  $F = a m$  also, but for more complicated cases such as a membrane subject to both residual stress and bending moments, it is nearly impossible to set up the equilibrium of forces. Therefore, it is an advantage that it is enough to calculate the potential and kinetic energy of such a membrane.

The kinetic energy  $E_{kin,m}$  is calculated from the integral over the kinetic energies of all infinitesimal elements of the membrane. The kinetic energy  $dE_{kin}$  of an



**Fig. 63** Vibrating spring loaded with a mass

Fig. 64 Vibrating membrane



infinitesimal volume element  $d_M dx dy$  of a membrane with thickness  $d_M$  and density  $\rho_M$  is given by (cf. Fig. 64):

$$dE_{kin} = \frac{\rho_M d_M dx dy}{2} v^2 = \frac{\rho_M d_M dx dy}{2} \left( \frac{\partial w}{\partial t} \right)^2. \tag{134}$$

The total kinetic energy is the integral over this equation:

$$E_{kin} = \iint \frac{\rho_M d_M}{2} \left( \frac{\partial w}{\partial t} \right)^2 dx dy. \tag{135}$$

The membrane deflection is approximated now with the following ansatz:

$$w(x, y, t) = w_0 u(x, y) \cos(\omega t + \varphi). \tag{136}$$

In this equation,  $u(x, y)$  denotes the part of the ansatz which is a function of the position and  $\cos(\omega t + \varphi)$  is the part which is a function of time. The kinetic energy in (135) is expressed now with the derivative of (136):

$$E_{kin} = \frac{\rho_M d_M}{2} w_0^2 \omega^2 \sin^2(\omega t + \varphi) \iint u^2 dx dy. \tag{137}$$

The kinetic energy is maximal when the sine function is equal to one, and this needs to be equal to the maximum of the potential energy  $V_{p,m}$ :

$$E_{kin,m} = \frac{\rho_M d_M}{2} w_0^2 \omega^2 \iint u^2 dx dy = V_{p,m}. \tag{138}$$

This equation is solved for  $\omega$  achieving the resonance frequency  $f$  of the membrane:

$$f = \frac{\omega}{2 \pi} = \frac{1}{2 \pi} \sqrt{\frac{2 V_{p,m}(u)}{\rho_M d_M w_0^2 \iint u^2 dx dy}}. \tag{139}$$

The maximum potential energy  $V_{p,m}$  can be calculated with the Ritz method (44 or 45 on page 35). The unknown shape  $u(x, y)$  of the vibrating membrane is

approximated with a reasonable ansatz containing unknown parameters. The unknown parameters are found then by calculating the minimum of the resonance frequency as a function of these parameters. Every deviation from the real shape of the vibrating membrane results in a higher frequency, and, therefore, the minimum approximates the real frequency [41]. With a suitable ansatz more than one relative minimum may be found. The least of these minima is an upper limit of the fundamental frequency and the larger ones correspond to higher-order resonances.

Typically, the static deflection shape  $u_1(x, y)$  of the membrane is a good ansatz for the deflection shape of the fundamental resonance frequency. Here, the shape of a thick membrane statically deflected by a constant pressure difference is assumed as described by (23) (page 29):

$$u_1(r) = \left(1 - \frac{r^2}{R_M^2}\right)^2. \quad (140)$$

With this ansatz, the only unknown parameter in (139) is the maximum deflection  $w_0$  of the vibrating membrane also called the amplitude of the vibration. When the potential energy is inserted into (139) from (44) or (45) (page 35) neglecting the term of stress due to straining of the neutral fiber, the vibration amplitude is canceling out. Thus, the resonance frequency is obtained directly from (139) and no further calculation of the minimum is required. For a circular membrane it is obtained:

$$f_1 = \sqrt{\frac{5}{3}} \frac{1}{\pi R_M \sqrt{\rho_M}} \sqrt{\frac{4}{3} \frac{E_M}{1 - \nu_M^2} \frac{d_M^2}{R_M^2} + \sigma_0}. \quad (141)$$

The fundamental frequency of a rectangular membrane is found with the ansatz, which was employed for the static deflection of square membranes also [(65) on page 45]:

$$u_1(x, y) = \left(1 - 4 \frac{x^2}{a_M^2}\right)^2 \left(1 - 4 \frac{y^2}{b_M^2}\right)^2 \quad (142)$$

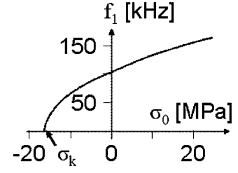
and yields:

$$f_{1,1} = \frac{\sqrt{3}}{\pi \sqrt{\rho_M}} \sqrt{\frac{1}{2} \frac{E_M}{1 - \nu_M^2} d_M^2 \left(\frac{7}{a_M^4} + \frac{4}{a_M^2 b_M^2} + \frac{7}{b_M^2}\right) + \left(\frac{\sigma_a}{a_M^2} + \frac{\sigma_b}{b_M^2}\right)}. \quad (143)$$

In these equations,  $\sigma_a$  and  $\sigma_b$  denote the stresses in the direction of the edges of the rectangular membrane with the lengths  $a_M$  and  $b_M$ , respectively. The stresses are assumed to be constant all over the membrane.

Figure 65 shows the fundamental frequency calculated with (141) as a function of membrane stress. A circular membrane with a radius and thickness of 0.5 mm and 5  $\mu\text{m}$ , respectively, and a density, Young's modulus, and Poisson's ratio of 1 kg/L,

**Fig. 65** Fundamental frequency of a circular membrane calculated with (141)



109 GPa, and 0.3, respectively, were assumed for this calculation. The figure and the equation show that the resonance frequency is decreasing with decreasing stress. At a certain compressive stress, the frequency becomes zero. This is the critical stress  $\sigma_k$  already known from (54) on page 41. Thus, (141) can be written in the form:

$$f_1 = \sqrt{\frac{5}{3}} \frac{1}{\pi R_M} \sqrt{\frac{\sigma_0 - \sigma_k}{\rho_M}} \quad \text{with} \quad \sigma_k = -\frac{4}{3} \frac{E_M}{1 - \nu_M^2} \frac{d_M^2}{R_M^2}. \quad (144)$$

Up to now, the medium surrounding the membrane has not been taken into account. Thus, the equations above are valid for membranes vibrating in vacuum only. If the membrane is in contact to a medium such as air or water, this fluid will vibrate together with the membrane and contribute to the moving mass. This is introduced by a so-called *additive mass*  $\beta_m$ . The following equation includes the effect of a fluid with density  $\rho_F$  which is in contact to one side of a circular membrane with density  $\rho_M$  [42]:

$$f_1 = \sqrt{\frac{5}{3}} \frac{1}{\pi R_M} \sqrt{\frac{\sigma_0 - \sigma_k}{\rho_M(1 + \beta_m)}} \quad \text{with} \quad \beta_m = \frac{2}{3} \frac{\rho_F}{\rho_M} \frac{R_M}{d_M}. \quad (145)$$

Most often the membrane will be in contact to the fluid on both sides, and, as a consequence, the additive mass needs to be doubled.

The effect of the additive mass was observed with micromembrane sensors as well. Figure 66a shows the measured frequency as a function of the density of a water solution. Potassium iodide (KI) was added stepwise enhancing the density with only little changes of the viscosity [39].

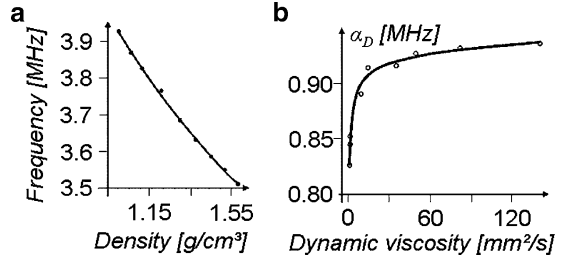
Up to now the damping of a vibrating membrane has not been taken into account here. The effect of the damping of a membrane (and any other vibrating body) is described by an exponential function. Thus, the maximum deflection is a function of time and the overall deflection becomes (cf. 136):

$$w(x, y, t) = A_0 e^{-\alpha_D t} u(x, y) \cos(\omega t + \varphi), \quad (146)$$

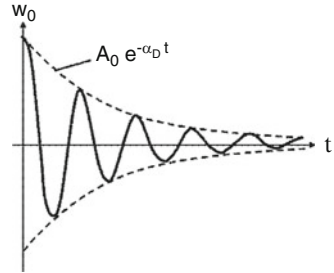
$$w_0 = A_0 e^{-\alpha_D t}, \quad (147)$$

where  $A_0$  and  $\alpha_D$  denote the vibration amplitude and the damping constant, respectively. Figure 67 shows the maximum deflection as a function of time according to the above equation.

**Fig. 66** (a) Measured resonance frequency of micromembrane sensors as a function of the density of the surrounding fluid and (b) damping coefficient as a function of the dynamic viscosity [39]



**Fig. 67** Deflection of a vibrating body as a function of time



Damping lowers the resonance frequency also. The angular frequency  $\omega$  of the damped vibration is described as a function of the undamped angular velocity  $\omega_0$  by the following equation:

$$2 \pi f = \omega = \sqrt{\omega_0^2 - \alpha_D^2}. \tag{148}$$

There are several reasons for damping such as the emission of sound and the friction in the surrounding fluid. The friction of the fluid is a function of its viscosity as shown in Fig. 66b, which displays the measured frequency of micromembrane sensors as a function of the viscosity enhanced by stepwise adding polyvinylmethylether (PVM) which has approximately the same density as water but a much larger viscosity. The effect of friction may be enhanced by the squeeze film effect also (cf. page 119ff).

The *damping constant due to the emission of sound* into a fluid on one side of a circular membrane is calculated by the following equation [42]:

$$\alpha_D = \frac{5}{36} \frac{\rho_F}{\rho_M c} \frac{R_M^2}{d_M (1 + \beta_m)^2} \frac{f_0^2}{4 \pi^2}. \tag{149}$$

In this equation,  $c$  denotes the velocity of sound in the fluid and  $f_0$  is the resonance frequency without damping due to sound emission. Often the membrane will be in contact to the fluid on both sides. As a consequence, both the additive mass needs to be doubled and the entire equation needs to be multiplied by two. This is illustrated

very well by an experiment with micromembrane sensors. Instead of a sensor with entrapped air as shown in Fig. 62a (page 88), a sensor with a frame open at the reverse side was used (cf. Fig. 68a). As a consequence, the micromembranes were emitting ultrasound to both sides and damping was larger. As shown in Fig. 68b, the vibrations of this sensor were damped out much quicker compared with Fig. 62d.

Damping of a vibrating element may be an advantage. For example, micromembrane sensors and the membranes of loud speakers of hearing aids need to emit enough sound for good performance. An element which emits sound also absorbs it very well. Thus microphones need to show large damping by sound emission also. Equation (149) shows that a larger damping can be achieved by increasing the diameter of a membrane and employing thin membranes with a small density.

The resonance frequencies of *vibrating beams* are calculated in a similar way as for membranes. Again there is a differential equation which needs to be solved with the required boundary conditions to obtain the frequencies. This differential equation is:

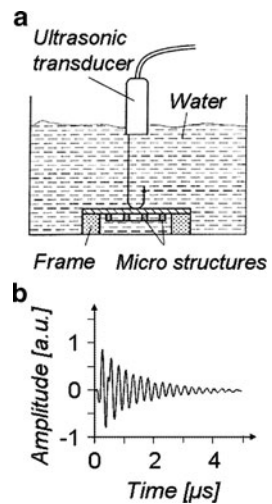
$$E_B I \frac{\partial^4 w}{\partial x^4} + \rho_B A_B \frac{\partial^2 w}{\partial t^2} = 0, \tag{150}$$

where  $E_B$ ,  $A_B$ ,  $\rho_B$ ,  $I$ , and  $x$  are Young's modulus, cross-sectional area, density, area moment of inertia of the beam, and the coordinate in the direction of the beam, respectively.

The solutions of this differential equation are:

$$w_i(x, t) = u_i(x) \cos(2 \pi f_i t + \varphi_i). \tag{151}$$

In this equation,  $w_i(x, t)$  describes the deflection of the beam as a function of the position  $x$  along the beam, the time  $t$ , and the parameter  $i$  which denotes the



**Fig. 68** Micromembrane sensor with an open frame (a) and signal measured with this sensor (b)

so-called mode of vibration. Every mode corresponds to a certain deflection shape  $u_i(x)$  which needs to fulfill both the boundary conditions and the differential equation. The resonance frequencies are the distinct values at which the differential equation is solved. The general solution of the differential equation includes all resonance frequencies and is the sum of all possible solutions:

$$w(x, t) = \sum_i A_i w_i(x, t) = \sum_i A_i u_i(x) \cos(2 \pi f_i t + \varphi_i). \quad (152)$$

The amplitudes  $A_i$  in this equation are a function of the excitation of the vibration.

The deflection shape solves the differential equation (150), if it is of the following form:

$$u_i(x) = a_1 \left[ \cosh\left(\frac{\lambda_i x}{L_B}\right) + a_2 \cos\left(\frac{\lambda_i x}{L_B}\right) - \sigma_i \sinh\left(\frac{\lambda_i x}{L_B}\right) + a_3 \sin\left(\frac{\lambda_i x}{L_B}\right) \right], \quad (153)$$

In this equation,  $a_1$  through  $a_3$ ,  $\sigma_i$ , and  $\lambda_i$  are coefficients which are to be chosen such that the boundary conditions are fulfilled. These boundary conditions are that the deflection  $u_i(x)$  of the beam and its first derivative are zero at  $x = 0$  where the beam is clamped and that the curvature of the beam (its second derivative, cf. page 6) is zero at its free end:

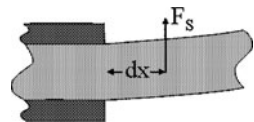
$$u_i(0) = 0 \Rightarrow a_2 = -1 \quad (154)$$

$$\left. \frac{\partial w}{\partial t} \right|_{x=0} = 0 \Rightarrow a_3 = \sigma_i \quad (155)$$

$$\left. \frac{\partial^2 w}{\partial x^2} \right|_{x=L_B} = 0 \Rightarrow \sigma_i = \frac{\cosh(\lambda_i) + \cos(\lambda_i)}{\sinh(\lambda_i) + \sin(\lambda_i)} = \frac{\sinh(\lambda_i) - \sin(\lambda_i)}{\cosh(\lambda_i) + \cos(\lambda_i)}. \quad (156)$$

A pair of shear forces  $F_S$  acting on a beam generates a bending moment (Fig. 69). If the axial distance at which the shear forces act is  $dx$ , the generated bending moment  $dM$  is:

$$dM = F_S dx \Rightarrow F_S = \frac{\partial M}{\partial x}. \quad (157)$$



**Fig. 69** Shear force generating a bending moment



The bending moment of a beam is described by (7) on page 7. Introducing this into the above equation, yields:

$$F_S = -E_B I \frac{\partial^3 w}{\partial x^3}. \tag{158}$$

Thus, as a consequence of the fact that no forces are acting on the free end of a beam, the third derivative of the deflection curve needs to be zero at the free end:

$$\left. \frac{\partial^3 w}{\partial x^3} \right|_{x=L_B} = 0 \Rightarrow \cos(\lambda_i) \cosh(\lambda_i) = -1. \tag{159}$$

The frequency parameters  $\lambda_i$  and  $\sigma_i$  calculated with these equations are shown in Table 7 with the precision of five digits. It is important to perform the calculations of the deflection shape with this accuracy, because otherwise the boundary conditions are not fulfilled.

$a_i$  is chosen such that the largest deflection which occurs along the beam is one. This way, the amplitude  $A_i$  in (152) corresponds to the largest deflection of the beam. As a consequence, the deflection shape of a beam clamped at one end is:

$$u_i(x) = \frac{1}{2} \left\{ \cosh\left(\frac{\lambda_i x}{L_B}\right) - \cos\left(\frac{\lambda_i x}{L_B}\right) - \sigma_i \left[ \sinh\left(\frac{\lambda_i x}{L_B}\right) - \sin\left(\frac{\lambda_i x}{L_B}\right) \right] \right\}. \tag{160}$$

When (151) and (160) are inserted into the differential equation (150), the following equation results, which yields the resonance frequencies of the beam:

$$\frac{\lambda_i^4}{L_B^4} - \frac{\rho A_B}{E_B I} (2 \pi f_i)^2 = 0 \Rightarrow f_i = \frac{\lambda_i^2}{2 \pi L_B^2} \sqrt{\frac{E_B I}{\rho_B A_B}}. \tag{161}$$

The deflection shape  $u_i(x)$  of a vibrating beam, 1 mm in length, is shown in Fig. 70a for the first three modes as calculated with (160) and Table 7.

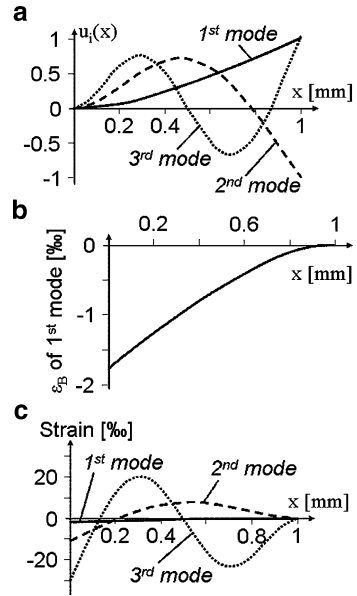
If a vibrating beam is to be employed as a sensor, strain gauges are a good way to measure the vibration amplitude and frequency. The resistance change as a function of the position along the beam can be calculated when the strain on the beam surface is known. The strain  $\varepsilon_B$  on the surface of the beam generated by bending is according to (5) on page 7:

$$\varepsilon_B = -\frac{d_B}{2} \frac{\partial^2 w}{\partial x^2}. \tag{162}$$

**Table 7** Frequency parameters of beams clamped at one end

Mode i	$\lambda_i$	$\sigma_i$
1	1.8751	0.7341
2	4.6941	1.0185
3	7.8548	0.9992

**Fig. 70** Deflection shape (a) and surface strain (b) and (c) of a vibrating beam clamped at one end



Inserting (152) and (160) yields:

$$\epsilon_B = -\frac{d_B}{2 L_B^2} \sum_i \left[ A_i \lambda_i^2 \frac{1}{2} \left\{ \cosh\left(\frac{\lambda_i x}{L_B}\right) + \cos\left(\frac{\lambda_i x}{L_B}\right) - \sinh\left(\frac{\lambda_i x}{L_B}\right) + \sin\left(\frac{\lambda_i x}{L_B}\right) \right\} \cos(2 \pi f_i t + \varphi_i) \right]. \quad (163)$$

In Fig. 70b, c, the strain  $\epsilon_B$  of the first three modes  $i$  is shown as calculated with this equation for the beam mentioned assuming a thickness of 100  $\mu\text{m}$  and an amplitude of 10  $\mu\text{m}$ . The resistance change of strain gauges mounted on top of the surface of the beam can be calculated by inserting (163) into (99) and (100) on page 68. If strain gauges are produced by doping of monocrystalline silicon, the resistance change is a function of crystal orientation also. The resistance changes then need to be calculated with (21) on page 22.

As both for strain gauges from silicon and from other materials, the resistance change is proportional to the strain, the largest effect may be expected from placing them in the near of the clamped end. As usual, strain gauges parallel and perpendicular to the beam may be combined compensating for temperature effects.

If the contribution of other modes than the desired one needs to be excluded from the measurement, an electronic band pass filter can be employed and/or the strain gauges are placed around the zeros of the strain of undesired modes. In Fig. 70c, this

would be at 0.2 and 0.5 mm for suppressing the second and third mode, respectively. It could be possible also to measure the strain near the clamping point and to subtract values proportional to the signal of strain gauges at the antinodes of undesired modes.

Often there is a *mass fixed at the free end of a vibrating beam* fixed at one end. For example, most acceleration sensors consist of a beam with a seismic mass at their end (cf. page 83). If the mass is approximated by a point mass  $m_0$  at the end of the beam and the mass of the beam is  $m_B$ , deflection shape  $u_1(x)$  and fundamental frequency  $f_1$  of this beam are calculated by:

$$u_1(x) = \frac{1}{2} \left( \left( 1 - \frac{x}{L_B} \right)^3 + 3 \frac{x}{L_B} - 1 \right), \tag{164}$$

$$f_1 = \frac{1}{2 \pi} \sqrt{\frac{3 E_B I}{L_B^3 (m_0 + 0.24 m_B)}}. \tag{165}$$

If a *vibrating beam is clamped at both ends* and only bending moments are in action as elastic forces (i.e., there is no residual stress), the same differential equation (150) applies as for a beam clamped only at one end. The only difference is that the boundary conditions are changed. As a consequence, the same equations are valid for the deflection shape (160) and resonance frequencies (161). However, the changed boundary conditions result in a changed factor  $a_1$  ( $=\frac{1}{2}$  in 160) and different frequency parameters  $\lambda_i$  and  $\sigma_i$ . They are derived from the boundary condition that both deflection and slope of the beam are zero where it is clamped. Thus, (154) and (155) need to be fulfilled not only at  $x = 0$  but also at  $x = L_B$ . Zero deflection at the end of the beam results in:

$$\sigma_i = \frac{\cos(\lambda_i) - \cosh(\lambda_i)}{\sin(\lambda_i) - \sinh(\lambda_i)} \tag{166}$$

and the condition of zero beam slope at its end yields:

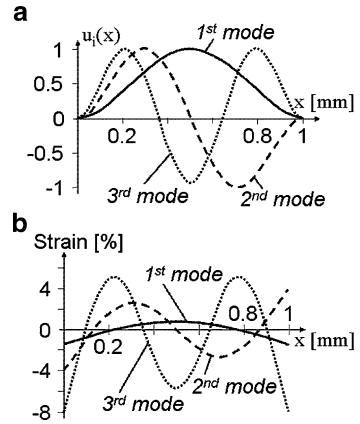
$$\cos(\lambda_i) \cosh(\lambda_i) = 1. \tag{167}$$

Table 8 shows the frequency parameters  $\lambda_i$  and  $\sigma_i$  calculated with these equations and Fig. 71a shows the deflection shape calculated with these parameters and (160).

**Table 8** Frequency parameters of beams clamped at both ends

Mode i	$\lambda_i$	$\sigma_i$
1	4.7300	0.9825
2	7.8532	1.000777
3	10.9956	0.999966

**Fig. 71** Deflection shape (a) and surface strain (b) of a vibrating beam without residual stress clamped at both ends



The factor  $\frac{1}{2}$  was replaced by 0.63 for the first mode, and 0.66 for the second and third mode, achieving a maximum deflection of 1.

The strain at the surface of the beam is calculated with (163) as in the case of a beam clamped only at one end, but with the frequency parameters from Table 8. The resistance change of strain gauges on the beam is obtained by inserting (163) into (99) and (100) on page 68. Figure 71b shows the strain on the surface of the beam. The resistance change of strain gauges is proportional to this. Again the largest effect is found in the near of the clamping points. If specific modes are to be selected, as in the case of a beam clamped only at one end, it is possible to place strain gauges at the antinodes or the nodes of the corresponding modes or to use frequency filters.

If a certain mode of vibration shall be excited, it is important to deflect it with the right resonance frequency. As shown later in this book, there are several possibilities to generate a beam deflection such as local heating (page 164f), electrostatic actuation (page 131f), and piezos (page 144f). The vibration at a certain mode is facilitated by generating beam bending at the antinodes or by generating a deflection at these positions.

Sometimes the vibrations of a *beam with a concentrated mass at its center* need to be known for designing. The deflection shape of the fundamental mode of this beam is described by:

$$u_1(x) = 4 \left( 3 \left( \frac{x}{L_B} \right)^2 - 4 \left( \frac{x}{L_B} \right)^3 \right). \tag{168}$$

This equation describes only half of the symmetrical deflection shape of the beam. The x-coordinate is starting at one clamping point. The fundamental resonance frequency of the beam is:

$$f_1 = \frac{4}{\pi} \sqrt{\frac{3 E_B I}{L_B^3 (m_0 + 0.37 m_B)}}. \tag{169}$$

All the equations shown above for a beam clamped at both ends do not include residual stress. However, *residual stress* has a significant *influence on resonance frequencies* and cannot be avoided, in general. It can be approximately taken into account by a factor multiplied to the resonance frequency without residual stress [43]:

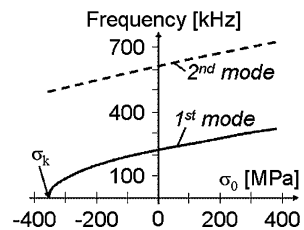
$$f_i(\sigma_0) = \sqrt{1 + \frac{\lambda_i^2}{\lambda_i^2} \frac{\sigma_0}{|\sigma_k|}} f_i(\sigma_0 = 0). \tag{170}$$

The frequency shift due to residual stress in a beam is shown in Fig. 72 as calculated with this equation. It is clearly seen that the fundamental frequency approaches zero when the compressive stress of the beam converges against the buckling load  $\sigma_k$ .

If the resonance frequency of a vibrating beam clamped at both ends shall be a measure of a force or a stress, respectively, it appears to be advantageous to employ the fundamental frequency of a beam near its critical stress  $\sigma_k$ , because the slope of the curve in Fig. 72 is largest there. If other parameters such as the mass of substances deposited on the beam shall be determined, it is desirable to avoid the effect of residual stress and a beam with more stress or vibrating at a higher mode may be preferred.

The *resonance frequencies* of all arrangements of beams with a mass fixed at their end which is much larger than the mass of the vibrating beam can easily be calculated when the force required for their deflection is known. Table 6 (page 80) shows the forces required for deflecting beams of several cases. The spring constant  $k$  of a certain case is the ratio of force  $F$  and deflection  $w_0$ , and the frequency can be calculated with (133) on page 89.

*Torsional vibrations* of microbeams occur easily, because the torque required to twist a beam by a certain angle  $\varphi$  (cf. Fig. 57 on page 78) is proportional to the third power of the linear dimensions [cf. (126) on page 79]. Thus, if torsional vibrations are undesired, this needs to be avoided by a suitable design. On the other hand,



**Fig. 72** Resonance frequencies of a vibrating beam as a function of residual stress calculated with (170) and (161)

torsional vibrations may be employed to measure small changes of a force or a torque.

Similar as in the case of transversal vibrations of beams [(152) on page 95] in general, the torsional angel  $\theta$  of a beam clamped only at one end is described by a sum over wave functions  $\theta_i$ , which need to fulfill both the boundary conditions and the differential equation:

$$\theta(x, t) = \sum_i A_i \theta_i(x, t) = \sum_i A_i u_i(x) \cos(2 \pi f_i t + \varphi_i). \tag{171}$$

In this equation,  $A_i$  denotes the amplitude,  $u_i(x)$  the deflection shape,  $f_i$  the resonance frequency, and  $\varphi_i$  a phase angle of the mode  $i$  of the torsional vibration. The boundary condition of a beam clamped only at one end is fulfilled, if

$$u_i(x) = \sin\left(\lambda_i \frac{x}{L_B}\right), \tag{172}$$

where  $\lambda_i$  is a solution of the following transcendental equation:

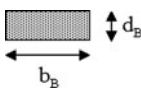

$$\cot(\lambda_i) = \frac{I_m}{\rho_B L_B I_t} \lambda_i. \tag{173}$$

In this equation,  $\rho_i$ ,  $L_B$ ,  $I_m$ , and  $I_t$  are the density, length, mass moment of inertia, and torsional constant (cf. Table 5 on page 78) of the beam, respectively. The mass moment of inertia is defined by the following integral:



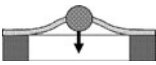

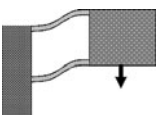

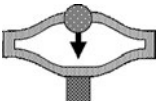
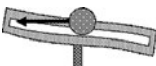
$$I_m = \int r^2 dm, \tag{174}$$

where  $r$  and  $dm$  are the distance from the axis of rotation and the infinitesimal mass element of the beam. The mass moments of inertia of beams often employed in microtechnique are shown in Table 9. For other cross-sections, the corresponding values are found in text books such as [32, 33].

**Table 9** Mass momentums of inertia of homogeneous beams rotating around their length axis as a function of their cross-section

Form of cross-section	Mass momentum of inertia
Rectangle 	$\frac{m_B}{12} (b_B^2 + d_B^2)$
Circle 	$\frac{m_B}{2} R_B^2$

**Table 10** Equations for the calculation of the resonance frequencies  $f$  of membranes and beams

Type	Equation
Circular membrane	$f_1 = \sqrt{\frac{5}{3}} \frac{1}{\pi R_M} \frac{1}{\sqrt{\rho_M}} \sqrt{\sigma_0 - \sigma_k}$ with $\sigma_k = -\frac{4}{3} \frac{E_M}{1 - \nu_M^2} \frac{d_M^2}{R_M^2}$
Rectangular membrane	$f_{1,1} = \frac{\sqrt{3}}{\pi \sqrt{\rho_M}} \sqrt{\frac{1}{2} \frac{E_M}{1 - \nu_M^2} d_M^2 \left( \frac{7}{a_M^4} + \frac{4}{a_M^2 b_M^2} + \frac{7}{b_M^4} \right) + \left( \frac{\sigma_a}{a_M^2} + \frac{\sigma_b}{b_M^2} \right)}$
Square membrane	$f_{1,1} = \frac{\sqrt{6}}{\pi a_M \sqrt{\rho_M}} \sqrt{\frac{\sigma_a + \sigma_b}{2} - \sigma_k}$ with $\sigma_k = -\frac{9}{2} \frac{E_M}{1 - \nu_M^2} \frac{d_M^2}{a_M^2}$
Beam clamped at both ends, $\lambda_i$ see Table 8 (page 98)	$f_1 = \frac{\lambda_1^2}{2 \pi L_B} \sqrt{\frac{E_B I}{L_B m_B}} \sqrt{1 - \frac{\lambda_1^2 \sigma_0}{\lambda_i^2 \sigma_k}}$ with $\sigma_k = -\frac{4\pi^2}{A_B} \frac{E_B I}{L_B^2}$
	$f_1 = \frac{\lambda_1^2}{2 \pi L_B} \sqrt{\frac{E_B I}{L_B m_B}}$ , $\lambda_i$ see Table 7 (page 96)
	$f_1 = \frac{\sqrt{3}}{2 \pi L_B} \sqrt{\frac{E_B I}{L_B (m_0 + 0.24 m_B)}}$
	$f_1 = \frac{4\sqrt{3}}{\pi L_B} \sqrt{\frac{E_B I}{L_B (m_0 + 0.37 m_B)}}$
	$f_1 = \frac{1}{2 \pi} \sqrt{\frac{E_B A_B}{m_0 L_B}}$
	$f_1 = \frac{1}{2 \pi} \sqrt{\frac{24 E_B I}{m_0 L_B^3}}$
	$f_1 = \frac{\lambda_i}{2 \pi d_B L_B} \sqrt{6 \frac{E_B}{1 + \nu_B} \frac{L_B I_t}{m_B}}$ , $\lambda_i$ from (173)
	$f_1 = \frac{1}{2 \pi} \sqrt{\frac{96 E_B I}{m_0 L_B^3}}$
	$f_1 = \frac{1}{2 \pi} \sqrt{\frac{(49 \pm 5) E_B I}{m_0 L_{Bc}^2 L_B}}$

For  $I$  and  $I_t$  see Table 3 (page 67) and Table 5 (page 78), respectively

Table 10 summarizes the resonance frequencies of circular membranes and beams with different boundary conditions.

## Exercises

### Problem 14

The fluid surrounding a membrane and the temperature have an effect on the fundamental frequency of a membrane.

- (a) Calculate the fundamental frequency of two circular membranes in vacuum and in water. One membrane is from silicon and the other one from polyimide. The thickness of the membrane is 25  $\mu\text{m}$  and the radius is 2 mm. The initial stress  $\sigma_0$  is the same for both membranes ( $\sigma_0 = 35 \text{ MPa}$  at  $20^\circ\text{C}$ ).
- (b) What is the fundamental frequency of the polyimide membrane, if the temperature is enhanced from 20 to  $95^\circ\text{C}$ ? The polyimide membrane is stretched over a frame from steel. Take into account the effect of the thermal strain only, and neglect the geometrical changes of thickness and radius, and the change of the density surrounding the membrane.
- (c) The density of the surrounding fluid is changed by the temperature change as well. This has an effect on the additive mass and in this way on the fundamental frequency. Calculate the changed fundamental frequency of the polyimide membrane in air and water, when the temperature is enhanced from 20 to  $95^\circ\text{C}$ . Take into account the change of the additive mass and the initial stress. Hint: The change of the density of a gas can be calculated with sufficient accuracy with the fundamental equation for gases which says that the density is proportional to the inverse of the absolute temperature [K].
- (d) Which meaning do the above effects calculated by you have for the designer of membranes in microsystems?

Young's modulus of silicon	160 GPa	Young's modulus of polyimide	1.66 GPa
Poisson's ratio of silicon	0.23	Poisson's ratio of polyimide	0.41
Density of air at $20^\circ\text{C}$	1.2 g/L	Density of water at $20^\circ\text{C}$	998 g/L
Temperature coefficient of the polyimide membrane	$45 \times 10^{-6} \text{ 1/K}$	Temperature coefficient of the frame of the membrane	$12 \times 10^{-6} \text{ 1/K}$
Density of water at $95^\circ\text{C}$	962 g/L	Gas constant $R_G$	8.314 J/(mol K)
Density of silicon	$2.3 \text{ g/cm}^3$	Density of polyimide	$1.43 \text{ g/cm}^3$
Temperature coefficient of silicon	$2.3 \times 10^{-6} \text{ 1/K}$		



### Problem 15

You have started a new job in a company producing sensors for the measurement of machine vibrations. The sensors are fabricated with silicon technology and consist of beams clamped at one end with strain gauges.

- (a) First you shall get to know the vibration sensors. You get some sensors from the production line on which the strain gauges are not yet implemented. You investigate the geometrical dimensions of the beams with an SEM (scanning electron microscope) and find out that the beams have a width of  $50\ \mu\text{m}$ , a thickness of  $25\ \mu\text{m}$ , and are  $605\ \mu\text{m}$  long.  
Please calculate the resonance frequencies of the first three modes.
- (b) By what factor would the thickness or length of the beam need to be altered achieving a fundamental frequency of  $5,820\ \text{Hz}$ ? What is the effect of the width of the beam?
- (c) Where would you place strain gauges if the sensor shall be sensitive to the resonance frequency of the three lowest modes?
- (d) At one customer of your company damages of the machine occur especially at a frequency which corresponds to the third mode of your beam. When these vibrations occur, the machine shall be switched off. Where would you place the strain gauges on the sensor described before to obtain a sensor which is especially sensitive to the third mode and insensitive to noise from the fundamental mode? Take the value from Fig. 70 on page 97.
- (e) The most recent development of your company is a vibration sensor consisting of several silicon beams arranged parallel to each other (cf. Fig. E16). At the end of each beam, there is an enlarged area with a gold layer serving as a seismic mass. The output signal of this sensor as a function of frequency is shown in Fig. E17. Obviously, it is an advantage of this sensor that a wide range of vibration frequencies can be recorded.

The gold covered ends of the beams show dimensions of  $200 \times 200\ \mu\text{m}^2$ .

The silicon of the beam and the end mass show the same thickness as the beams of part (a), but the beam is  $2.5\ \text{mm}$  long. What mass is needed at the end of the beam to obtain a resonance frequency of  $3\ \text{kHz}$ ? Calculate the thickness of the gold layer.

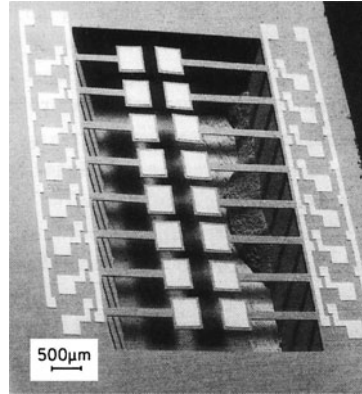
Hint: Assume for your calculation that the gold covered end would be concentrated as a point mass at the end of the beam.

- (f) What is the advantage of the gold covered ends compared with beams without any extrastructure at their end but the same resonance frequency?

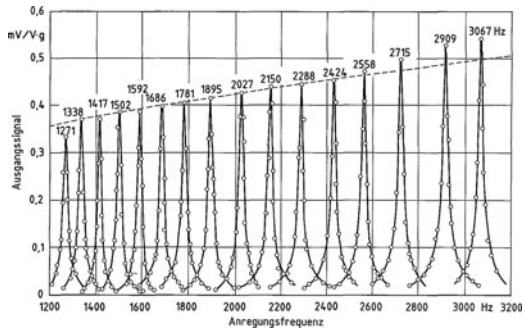
Density of silicon	$2.32\ \text{g/cm}^3$	Young's modulus of silicon	$160\ \text{GPa}$
Density of gold	$19.3\ \text{g/cm}^3$		

Hint: Assume for your calculations that the sensor would vibrate in vacuum and that the surrounding medium would show no effect on the vibrations.

**Fig. E16** Vibration sensor with several beams of different lengths and with gold covered seismic masses at their ends [44]



**Fig. E17** Output signal of the vibration sensor [44] from Fig. E16



### Problem 16

By microgravimetry the mass of very small particles and molecules can be measured. The change of the resonance frequency of membranes or beams as a function of the mass of these structures is employed to measure small masses in microgravimetry. A selective layer on the microstructure allows the binding of a certain kind of molecules only. This allows the analysis of DNA, RNA, and proteins. The performance of such microscales is demonstrated especially by so-called electronic noses: Even the binding of small molecules on the surface of a vibrating membrane or a vibration beam results in a change of the resonance frequency, and, therefore, can be employed to recognize the presence of a certain smell.

The microscale manufactured by you consists of a silicon beam clamped at one end. With a piezo, the beam can be excited to vibrations. The beam shows a length of 200 µm, a width of 50 µm, and a thickness of 5 µm. The free end of the beam is covered on both sides with proteins on an area of 50 × 50 µm<sup>2</sup>. On these proteins, only the macromolecules of the type A can adhere.

- (a) Please calculate the fundamental resonance frequency of the beam without any molecules on it.
- (b) At the beginning of your measurement, you perform a reference measurement. You obtain the resonance frequency calculated at part (a) of this exercise. Now, you conduct an air stream with the macromolecules of interest over the beam and measure the resonance frequency once more. Assume that the area on both sides of the beam becomes covered with a 100-nm thick layer with a density of  $1.2 \text{ g/cm}^3$ .  
What resonance frequency is measured now?

Hint:

- Assume that the molecules are distributed homogeneously on both sides of the beam and all remain to be adhered.
  - Assume for the calculation that the additional mass of the macromolecules is attached to the end of the beam.
- (c) How much does the resonance frequency change, when a dust particle with a mass of  $10 \text{ }\mu\text{g}$  would adhere at the end of the beam?

Hint: Assume for all problems of the exercise that you may use the equations of beam vibrations in vacuum. (In a real calculation, the effect of the surrounding air with its damping and additive mass would need to be taken into account.)

Young's modulus of silicon	160 GPa	Density of silicon	$2.32 \text{ g/cm}^3$
----------------------------	---------	--------------------	-----------------------

# Capillaries

Microfluidic is one of the most promising fields of microtechnique. A variety of microfluidic devices has been developed such as micropumps (page 229), microvalves (page 203), pressure sensors (page 275), flow sensors (page 289), and analysis systems. The most important basic element of microfluidics is the capillary. It is needed to link components and to direct the flow of a fluid. The notion “fluid” includes both gasses and liquids. Besides this, capillaries can be employed for the separation of different ingredients.

Capillaries are made of nearly every rigid material used in microtechnique. They typically show diameters between 10  $\mu\text{m}$  and some millimeters. It is possible to fabricate capillaries with a smaller diameter but the risk that they are blocked by a small particle rises quickly with decreasing diameter, and, therefore, most capillaries are wider than 100  $\mu\text{m}$ . The total length of a capillary is limited to several centimeters by the space available on a single device. On the other hand, a capillary shorter than approximately 100  $\mu\text{m}$  normally is not very useful.

Capillaries can be fabricated as a concave edge which is open to the environment. However, the evaporation rate of water is on the order of some millimeters per day. One millimeter per day is the same as 42  $\mu\text{m}/\text{h}$ . Thus, a capillary may be emptied by evaporation within a short time, and the concentration of an analyte in the liquid in a capillary may change significantly within seconds. That is why most capillaries need to be designed as closed channels avoiding evaporation.

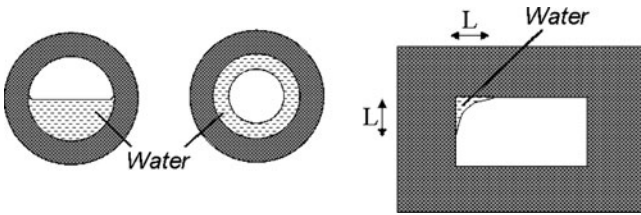
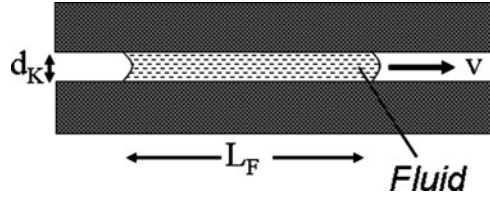
A certain pressure difference  $\Delta p_R$  is necessary to move a fluid inside of a capillary, because friction needs to be overcome. The pressure difference to move a fluid with length  $L_F$  and dynamical viscosity  $\eta$  in a capillary with the velocity  $v$  is given by the *Hagen Poiseuille equation* (cf. Fig. 73):

$$\Delta p_R = -32 \frac{\eta L_F}{D_h^2} v. \quad (175)$$

The quantity  $D_h$  denotes the so-called hydraulic diameter of the capillary. It expresses the influence of the geometrical shape of the cross-section of the capillary on the flow. The definition of the hydraulic diameter is:

$$D_h = \frac{4A_F}{U_w}, \quad (176)$$

**Fig. 73** Fluid in a capillary driven by a pressure drop



**Fig. 74** Partly filled capillaries. *Left:* Half filled macroscopic and microscopic pipes, respectively. *Right:* Water in a concave edge

where  $A_F$  is the area of the cross-section and  $U_w$  is the wetted circumference of the capillary. For a circular cross-section, the hydraulic diameter is the same as the real diameter.

The hydraulic diameter can also be employed in unusual cases such as a macroscopic pipe which is filled only partly. In a macroscopic pipe, e.g., the lower half could be filled with water (cf. Fig. 74 left). Thus, half of the cross-section and half of the circumference need to be introduced into (176). In a microscopic pipe, capillary forces could (hypothetically) cause the water to be attached to the walls. As a consequence, the complete circumference and half of the cross-sectional area would be used in (176).

The flow in a capillary formed by a concave edge can also be calculated with (176) (cf. Fig. 74 right). If the length of the liquid in touch with the wall is designated by  $L$  and its cross-sectional area is estimated to  $A = (3/8)L^2$ , the hydraulic diameter according to (176) is calculated as  $D_h = (3/4)L$ .

If a pressure difference  $\Delta p$  over a capillary is given or a liquid flow is driven by known capillary forces (cf. page 114f), the flow velocity  $v$  can be calculated from (175):

$$v = -\frac{1}{32} \frac{D_h^2}{\eta L_F} \Delta p. \tag{177}$$

From the mean flow velocity  $v$  from (177), the volume flow  $\Phi_F$  is calculated by multiplying with the cross-section  $A$  of the channel:

$$\Phi_F = v A = -\frac{1}{32} \frac{D_h^2 A}{\eta L_F} \Delta p =: \frac{\Delta p}{R_{\Pi}}. \tag{178}$$

The quantity  $R_{fl}$  in the above equation is the so-called flow resistance of the channel. Table 11 shows the hydraulic diameter and flow resistance of microchannels which are often used in microtechnique.

For a short capillary, (177) yields unrealistically high velocities. This is due to the fact that the inertial force needed for accelerating the liquid has not yet been taken into account. If a liquid is taken from a resting reservoir into a capillary (cf. Fig. 75), the force  $dF$  is required to achieve the acceleration  $a_F$  bringing the mass  $dm$  to the velocity  $v$  in the time  $dt$ :

$$dF = a_F dm = \frac{dv dm}{dt}. \quad (179)$$

The mass  $dm$  equals the density  $\rho_F$  of the liquid times the volume, and the volume is the cross-section  $A$  of the capillary times the length  $dx$  the liquid moves in the time  $dt$ :

$$dF = \frac{dv dm}{dt} = \frac{\rho_F A_F dx}{dt} dv = \rho_F A_F v dv. \quad (180)$$

Integrating on both sides and taking into account that the pressure is decreasing in the direction of the velocity yields:

$$\int dF = F = \rho_F A_F \int v dv = \rho_F A_F \frac{v^2}{2} \Rightarrow \Delta p = -\frac{\rho_F}{2} v^2. \quad (181)$$




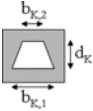

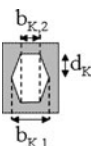

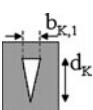
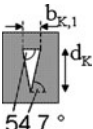
To the pressure drop  $\Delta p$  which is necessary to accelerate the liquid, the pressure drop needed to overcome the friction (175) has to be added:

$$\Delta p = -\frac{\rho_F}{2} v^2 - 32 \frac{\eta L_F}{D_h^2} v \Rightarrow v = -32 \frac{\eta L_F}{\rho_F D_h^2} + \sqrt{\left(32 \frac{\eta L_F}{\rho_F D_h^2}\right)^2 - \frac{2 \Delta p}{\rho_F}}. \quad (182)$$

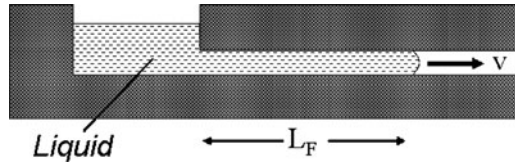
As a consequence of this equation, the flow into a capillary driven by a constant pressure drop or by constant capillary forces will start with a velocity which is a function of the force required for acceleration and then reduces as a result of the friction in a capillary with growing wetted length. Figure 76 shows the velocity calculated with (177) and (182) as a function of the length water (with dynamic viscosity  $\eta = 1$  mPa s, density  $\rho_F = 1,000$  kg/m<sup>3</sup>) has penetrated into a capillary with a square cross-section of  $100 \times 100 \mu\text{m}^2$  driven by a pressure drop of  $-10$  kPa.

Also important for the calculation of flow velocities in capillaries is the *Bernoulli equation*. If the cross-section of a capillary is reducing, the mean flow velocity  $v$  of an incompressible liquid needs to increase there, because the volume flow has to be constant everywhere along the capillary (Fig. 77). As described

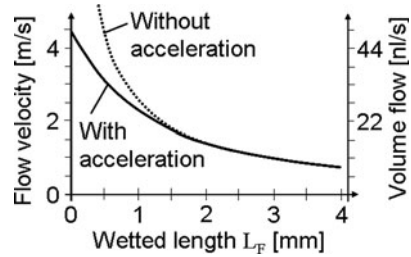
**Table 11** Hydraulic diameter and flow resistance of microchannels often used in microtechnique as calculated from (176) and (178)

	Hydraulic diameter $D_h$	Flow resistance $R_{fl}$
	$2 R_{Ka}$	$\frac{8 \eta L_K}{\pi R_{Ka}^4}$
	$2 \frac{b_K d_K}{b_K + d_K}$	$8 \eta L_K \frac{(b_K + d_K)^2}{b_K^3 d_K^3}$
	$b_K$	$\frac{32 \eta L_K}{b_K^4}$
	$\frac{2 d_K (b_{K,1} + b_{K,2})}{b_{K,1} + b_{K,2} + \sqrt{4 d_K^2 + (b_{K,1} - b_{K,2})^2}}$	$16 \eta L_K \frac{\left( b_{K,1} + b_{K,2} + \sqrt{4 d_K^2 + (b_{K,1} - b_{K,2})^2} \right)^2}{(b_{K,1} + b_{K,2})^3 d_K^3}$
	$2 d_K \frac{2 b_{K,1} - 1.416 d_K}{2 b_{K,1} + 1.035 d_K}$	$16 \eta L_K \frac{(2 b_{K,1} + 1.035 d_K)^2}{(2 b_{K,1} - 1.416 d_K)^3 d_K^3}$
	$\frac{4 (b_{K,1} + b_{K,2}) d_K}{2 b_{K,2} + 2 \sqrt{4 d_K^2 + (b_{K,1} - b_{K,2})^2}}$	$8 \eta L_K \frac{\left( b_{K,2} + \sqrt{4 d_K^2 + (b_{K,1} - b_{K,2})^2} \right)^2}{(b_{K,1} + b_{K,2})^3 d_K^3}$
	$2 d_K \frac{2 b_{K,1} - 1.416 d_K}{b_{K,2} + 2.451 d_K}$	$8 \eta L_K \frac{(b_{K,2} + 2.451 d_K)^2}{(2 b_{K,1} - 1.416 d_K)^3 d_K^3}$
	$\frac{2 d_K b_{K,1}}{b_{K,1} + \sqrt{4 d_K^2 + b_{K,1}^2}}$	$16 \eta L_K \frac{\left( b_{K,1} + \sqrt{4 d_K^2 + b_{K,1}^2} \right)^2}{b_{K,1}^3 d_K^3}$
	$0.478 b_{K,1} = 0.548 d_K$	$\frac{338.7 \eta L_K}{b_{K,1}^4} = \frac{84.25 \eta L_K}{d_K^4}$

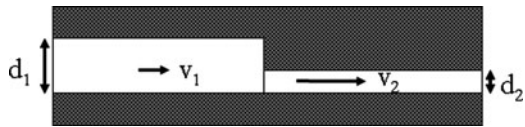
**Fig. 75** Fluid entering a capillary driven by a pressure difference



**Fig. 76** Flow velocity in a capillary as a function of the wetted length calculated with (177) and (182)



**Fig. 77** Capillary with changing cross-section



before, the velocity increase is combined with a pressure drop which provides the force required for acceleration:

$$p + \frac{\rho_F}{2} v^2 = p_0 = \text{const.} \tag{183}$$

The first term in this equation is the static pressure which can be measured with a pressure sensor perpendicular to the flow direction. The second term is the so-called dynamic pressure which corresponds to the kinetic energy of the flow. The Bernoulli equation says that the sum of static and dynamic pressure needs to be constant.

If the fluid is a gas, it is compressible and the calculation becomes more complex, because the density of the fluid increases with the pressure and, besides this, the temperature  $T$  may be a function of the flow velocity also. If the flow velocity is so small that temperature changes are equalized at once with the ambient, the flow is called isothermal and the pressure of a gas with mol mass  $m_{\text{mol}}$ , flow velocity  $v$ , and pressure at rest  $p_0$  is:

$$p = p_0 e^{-(m_{\text{mol}} v^2 / (2 R_G T))}, \tag{184}$$

where  $R_G = 8.31 \text{ J}/(\text{mol K})$  is the gas constant.



If the flow of a gas is so fast that there is not enough time for temperature equalization, it is called adiabatic, and pressure and temperature are calculated with the following equations:

$$p = p_0 \left( \frac{1}{2} \left[ \frac{C_p}{C_v} - 1 \right] \frac{v^2}{c^2} + 1 \right)^{C_p/(C_p+C_v)}, \quad (185)$$

$$T = T_0 \left( \frac{1}{2} \left[ \frac{C_p}{C_v} - 1 \right] \frac{v^2}{c^2} + 1 \right). \quad (186)$$

In these equations,  $C_p$ ,  $C_v$ , and  $c$  denote the heat capacity at constant pressure, constant volume, and the velocity of sound, respectively.

Figure 78 shows pressure and temperature as a function of flow velocity as calculated with (183), (184), (185), and (186). Incompressible and compressible isothermal flow show a significant difference only at velocities higher than 7 m/s. Such high velocities rarely occur in microtechnique. The heat capacities in microtechnique are very small, and, therefore, most flows may be considered as isothermal. Thus, in most cases, the pressure drop may be calculated with the Bernoulli equation (183). The temperature rise in adiabatic flows shows a significant effect only at very high velocities, and, therefore, not often is important in microtechnique.

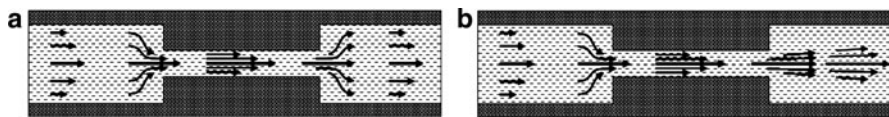
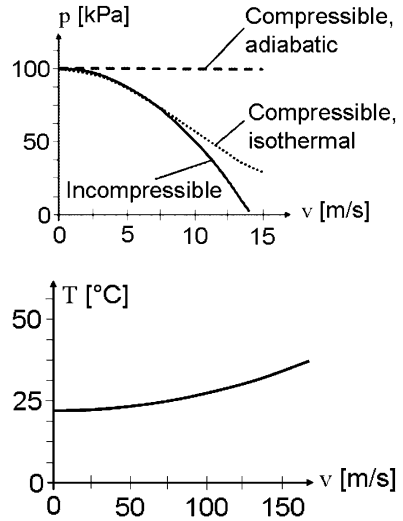
Where there is a constriction in a flow channel, the mean flow velocity needs to become larger (see above). As a consequence, the static pressure is reduced in the constriction. If the flow is slow, the mean velocity behind the constriction will be the same as in front of it (cf. Fig. 79a) and the static pressure behind the constriction is the one in front of it reduced by the pressure drop due to the Hagen Poiseuille equation [(175), page 107]. However, if the flow is quick, the flow will emanate from the end of the constriction as a jet. This jet may lose its kinetic energy due to friction (Fig. 79b). The energy which is absorbed by friction is lost and results in an additional pressure drop over the constriction. This pressure lost is a function of the flow velocity and the geometry at the entrance and the exit of the channel. Therefore, a quantity  $\chi$  is defined which describes the amount of this effect. The pressure drop due to the change of the mean velocity  $v$  is calculated with (183). Together with the Hagen Poiseuille equation it is obtained:

$$\Delta p = -\chi \frac{\rho_F}{2} (v_1^2 - v_2^2) - 32 \frac{\eta L_F}{D_h^2} v_2. \quad (187)$$

In the above equation,  $v_1$  and  $v_2$  are the mean flow velocities inside and outside the constriction, respectively. The volume flow  $\Phi_F$  is the product of mean flow velocity and cross-sectional area  $A$ , resulting in:

$$\Delta p = -\chi \frac{\rho_F}{2} \left( \frac{1}{A_1^2} - \frac{1}{A_2^2} \right) \Phi_F^2 - 32 \frac{\eta L_F}{D_h^2} \frac{\Phi_F}{A_2}. \quad (188)$$

**Fig. 78** Static pressure for incompressible, adiabatic, and isothermal flow and temperature for adiabatic flow as a function of the flow velocity calculated with (183–186)



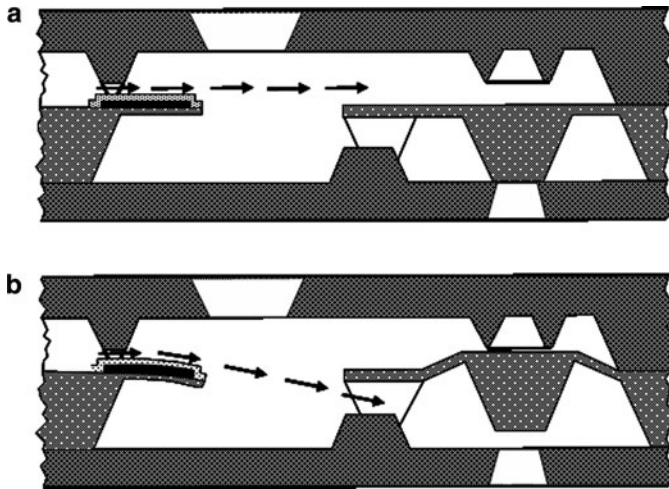
**Fig. 79** Slow (a) and quick (b) flow through a constriction

The second term in the above equation is the Hagen Poiseuille equation and the first term is called the Torricelli equation which describes the flow through an aperture.

A consequence of the Bernoulli effect is the *Coanda effect*. The Coanda effect is known from everyday experience: If coffee or tea shall be filled into a cup, it is spilled because the liquid stream is deflected from the expected way towards the can. Due to the Bernoulli effect, the pressure in the near of the liquid flow is decreased. This is compensated by air flowing towards the liquid stream. The wall of a rigid body (e.g., the can) in the near of the stream does not allow for this air flow, and, therefore, the liquid stream is deflected towards the wall.

This effect can be employed in microtechnique by steering a stream with a moving wall. Figure 80a shows a microvalve in which an air stream is emanating out of a nozzle over a bimetal stripe and closing the outlet of the valve with the pressure rise generated over the membrane [45]. If the bimetal stripe is heated, it is deflecting down and the air stream is directed below the membrane opening the valve (Fig. 80b).

Another application of the Coanda effect is shown in Fig. 81 [46]. An air stream emanating from a nozzle sticks to one of two walls and can be moved to the other



**Fig. 80** Microvalve employing the Coanda effect (reproduced from [45])

one with a short pressure pulse. Thus, the pressure at the two outlets of this device can be switched back and forth.

There are several ways how a fluid in a capillary can be moved. The most important ones are a pressure drop, capillary forces, electro osmosis, electrophoresis, dielectrophoresis, and capacitive forces. These are described in the following.

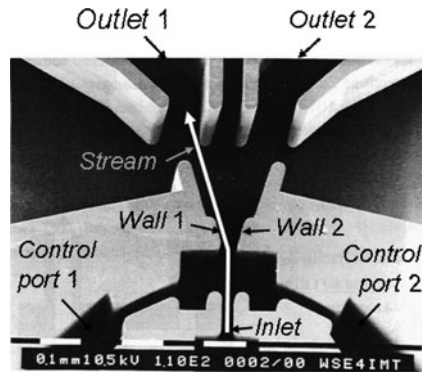
*Capillary forces* are the result of polarized atoms or molecules in the fluid and the rigid wall of the capillary attracting or repelling each other by electrostatic forces. Every fluid is more or less attracted into or repelled out of a capillary. As the force is generated at the interface between fluid and capillary wall, the circumference of the fluid is leading the flow and the volume is following due to attractive intermolecular forces inside of the fluid as shown in Fig. 82a, b. This is opposite to the situation of a fluid driven by a pressure drop (cf. Fig. 73 on page 108 and Fig. 75).

Capillary forces are acting on both gasses and liquids. However, the effect is more pronounced with liquids and gasses show a minor effect. When the capillary is filled completely or does not have a link to a reservoir outside the capillary, the intermolecular forces act in both directions in the same way. Thus, no effect on the fluid is observed any more.

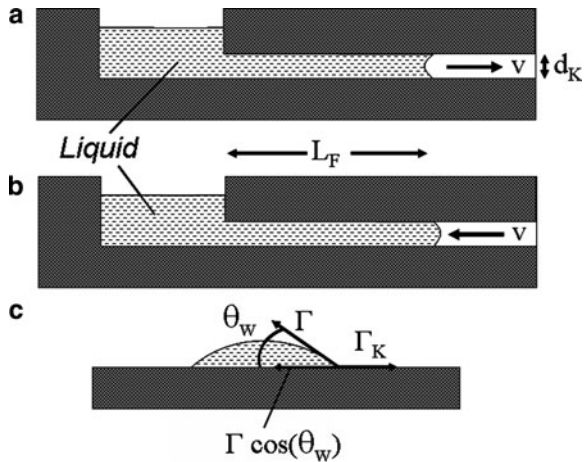
The strength of capillary forces is a function of the surface tension  $\Gamma$  of the fluid and the wetting angle  $\theta_w$  which is adopted by a liquid in contact with the wall material and another liquid or gas.

Figure 82c shows a droplet of a liquid in contact to a wall. At the rim of the droplet, an equilibrium is established between the forces  $\Gamma_K$  per circumference length which try to wet the wall on the one hand, and on the other hand the component of the surface tension  $\Gamma$  which does not want to get into contact to the surrounding gas and binding the molecules of the liquid together:

**Fig. 81** Fluidic beam element applied as a pneumatic flow control [46]. (Courtesy of Karlsruhe Institute of Technology, KIT)



**Fig. 82** A liquid attracted into (a) and repelled off (b) by capillary forces, and (c) the equilibrium of forces at the circumference of a liquid on a rigid body



$$\Gamma_K = \Gamma \cos(\theta_w). \tag{189}$$

This so-called “capillary force” is not really a force but a force per circumference length and the surface tension is not a tension either. The real force  $\Gamma_{Kap}$  attracting the liquid into the capillary or repelling it out of it is obtained by multiplying (189) with the circumference length  $U_w$ :

$$\Gamma_{Kap} = U_w \Gamma \cos(\theta_w). \tag{190}$$

From this equation, the capillary pressure ( $\Delta p_K =$  force per cross-sectional area  $A_F$ ) can be calculated which corresponds to the pressure drop which would result in

the same effect as the capillary force. It needs to be taken into account that the pressure drop has the opposite direction than the force:

$$\Delta p_K = -\frac{U_w}{A_F} \Gamma \cos(\theta_w). \quad (191)$$

This equation shows that the capillary pressure is negative (i.e., the liquid is attracted into the capillary), if the wetting angle  $\theta_w$  is smaller than  $90^\circ$  and that a liquid is repelled from a capillary, if the wetting angle is larger than  $90^\circ$  (i.e., the cosine is negative).

In general, the total pressure driving a liquid is the sum of the pressure drop  $\Delta p_p$  over the capillary and the capillary pressure  $\Delta p_K$ . This sum is equal to the pressure  $\Delta p_R$  required to overcome the friction [cf. (175) on page 107], if friction is much larger than the inertia force needed for acceleration:

$$\begin{aligned} \Delta p &= \Delta p_p + \Delta p_K = \Delta p_p - \frac{U_w}{A_F} \Gamma \cos(\theta_w) = -32 \frac{\eta L_F}{D_h^2} v = -2 \eta L_F \frac{U_w^2}{A_F^2} v \\ \Rightarrow v &= \frac{1}{2 \eta L_F} \frac{A_F^2}{U_w^2} \left( -\Delta p_p + \frac{U_w}{A_F} \Gamma \cos(\theta_w) \right) \end{aligned} \quad (192)$$

If a gap with width  $b_K$  and height  $d_K$  is filled with a liquid, the cross-sectional area of this capillary is  $A_F = b_K d_K$  and its wetted circumference is  $U_w = 2 b_K$  (cf. Fig. 83). Inserting this into (192), yields the flow velocity of a liquid in a gap:

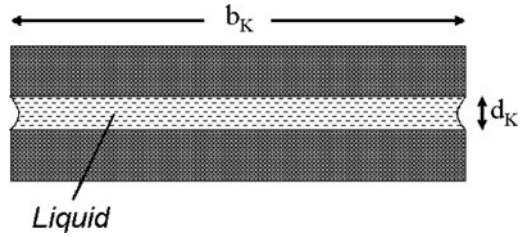
$$v = \frac{d_K^2}{8 \eta L_F} \left( -\Delta p_p + \frac{2}{d_K} \Gamma \cos(\theta_w) \right). \quad (193)$$

If the capillary is rectangular and width and height are denoted by  $b_K$  and  $d_K$ , respectively, the circumference is  $U_w = 2 b_K + 2 d_K$ , yielding:

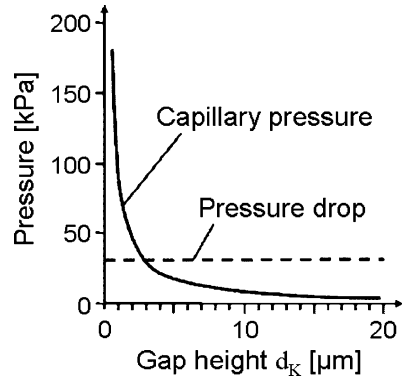
$$v = \frac{1}{8 \eta L_F} \frac{d_K^2 b_K^2}{d_K + b_K} \left( -\Delta p_p + 2 \frac{d_K + b_K}{d_K b_K} \Gamma \cos(\theta_w) \right). \quad (194)$$

The flow velocity in a gap as described by (193) shows two terms in the parenthesis: the pressure difference and the capillary pressure. The capillary pressure is inversely proportional to the height of the capillary, while a pressure drop from the outside is not a function of capillary size. Therefore, capillary pressure becomes larger than any given pressure when the capillary gap is reduced enough. Figure 84 shows both a pressure drop of 30 kPa and capillary pressure of water with a surface tension of  $\Gamma = 50$  mN/m and a wetting angle of  $\theta_w = 15^\circ$  calculated with (191) as a function of the gap height. This explains why capillary forces are not important in the macroscopic world, while they may play a dominating role in microtechnique.

**Fig. 83** Cross-section of a gap filled with a liquid. The flow direction is normal to the drawing area



**Fig. 84** Pressure drop and capillary pressure as a function of the height of a gap



If the height of a gap or the cross-section of a capillary increases, the capillary force is reduced significantly. The velocity of a flow driven by capillary force ( $\Delta p_p = 0$ ) in a gap as calculated with (193) at the end of the gap with height  $d_K$  in Fig. 85:

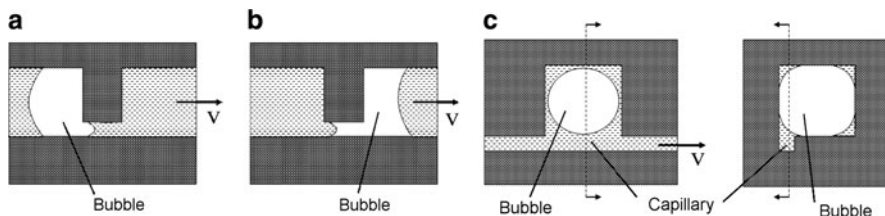
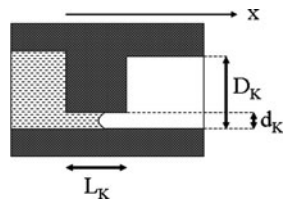
$$v = \frac{\Gamma \cos(\theta_w) d_K}{4 \eta L_K}. \tag{195}$$

Behind the lower gap, the capillary force is reduced, because the height  $D_K$  of the higher gap needs to be used in the term of the capillary force in the parentheses in (193), while the friction is still due to the lower gap and  $d_K$  needs to be used in the term before the parentheses resulting in:

$$v = \frac{\Gamma \cos(\theta_w) d_K}{4 \eta L_K} \frac{d_K}{D_K}. \tag{196}$$

Thus, the sudden widening of the gap reduces the flow velocity by the ratio of the two gap heights. As a consequence, the capillary flow is nearly stopped. This can be employed to direct a capillary flow. For example, adhesive used to bond two parts can be kept away from certain positions by a wide *stop groove* and the adhesive dries before it overcomes the groove with the reduced flow velocity.

**Fig. 85** Capillary flow arriving at a stop groove



**Fig. 86** Bubble blocking a constriction by (a) attractive and (b) repulsive capillary forces. (c) Bubble catching capillary part shown as a cross-section along the capillary (left) and perpendicular to it

A gas bubble in a liquid flowing through a capillary has nearly no effect on the velocity, if the cross-section of the capillary does not change. However, if the bubble arrives at a constriction, the capillary force needs to be overcome which either pulls the liquid into the narrower part of the capillary or rejects it from it (cf. Fig. 86a, b). Therefore, a liquid flow driven by a pressure drop over the capillary may be blocked when the bubble arrives at a constriction, if the driving pressure is not larger than the capillary pressure calculated with (191).

Obviously, a way to avoid problems with bubbles blocking capillaries is to design no constrictions. However, a lot of applications are based on a changing cross-section, e.g., microvalves. Thus, it is interesting to know that this problem can be circumvented. If the capillary forces are attractive, bubbles can be caught in a part of the capillary which shows an enlarged cross-section (cf. Fig. 86c). Bubbles are pushed into the larger cross-section by the liquid attracted into the smaller one. If capillary forces are repulsive, a system of smaller capillaries with a large enough volume is required next to the capillary.

Capillary forces are a function of temperature because surface tension  $\Gamma$  is a function of temperature which enters into (191). For most liquids, the interrelationship between surface tension and temperature is a linear function:

$$\Gamma = k_s(T_K - T), \quad (197)$$

where  $T_K$  and  $k_s$  are constants (for water  $T_K = 710$  K). For most liquids which boil below 200–300°C,  $k_s$  can be calculated as:

$$k_s = 2.1 \times 10^{-7} \left( \frac{m_{\text{mol}}}{\rho_F} \right)^{-2/3}, \quad (198)$$

where  $m_{\text{mol}}$  and  $\rho_F$  denote the mass of 1 mol of the liquid and its density, respectively. For water, the surface tension and the capillary force vary by approximately 20% between freezing and boiling point. This can be employed to build up a capillary motor. Heaters can be used to reduce the capillary force at one end and move a droplet, this way. Difficulties of this motor may be that the liquid may evaporate or get lost by diffusion through the walls of the capillary.

A special application of repulsive capillary forces is the so-called *lotus effect*. Water is repelled from the surface of the leaf of a lotus plant because the leaf is hydrophobic and has a microstructure on its surface. On a hydrophilic surface, water penetrates into every microcavity. Since the friction in such cavities is very large and water molecules are bonded to each other by intermolecular forces, water droplets on a hydrophilic surface with a microstructure are hindered from moving. The opposite happens on a hydrophobic surface. The water is repelled from the cavities on the surface and only very few and small contact areas remain between water and the surface. Therefore, the friction of water on such a surface is much reduced and does not adhere on the surface. Hydrophobic surfaces with a microstructure are also called *super hydrophobic surfaces*.

The lotus effect was first discovered on plants and insects in the 1970s and then employed for the development of foils, paints, and clothes which repel water and dirt.

If one wall of a capillary is moved in normal direction, the fluid inside is pressed out or sucked into the capillary. The fluid entering or ejected from the capillary is subject to the friction force calculated with (175) on page 107. As a result, a certain force is required to move the wall. This is the so-called *squeeze film effect* which is well known from the difficulty to lift the lid of a pot lying on a flat surface, e.g., the kitchen desk. This effect is also one of the reasons why a rotating coin needs such a long time to fall down onto a desk.

In general, the calculation of the squeeze film effect is very complex, because a lot of phenomena have to be taken into account. Here, an equation is derived for the capillary formed by a rectangular wall open to all sides (cf. Fig. 87) with the help of simplifications which, in general, are not fulfilled but allow calculating at least the correct order of magnitude of the effect. It is assumed that the fluid in the capillary moves only in the direction of the width of the capillary. This is approximately true, if the wall is much longer than wide. It is assumed further that the fluid is incompressible and that pressure changes due to the Bernoulli equation [(183) on page 111] are negligible. Besides this, deformations of the wall due to the pressure difference and the movement are not taken into account.

The volume  $dV$  which is squeezed out of the capillary between its center and the position  $x$  after the wall is moved by  $dz$  is given by:

$$dV = L_K x dz, \quad (199)$$



where  $L_K$  denotes the length of the wall. From this, it is easily derived that the volume change  $\dot{V}$  per time unit as a function of the velocity  $v_w$  of the wall is:

$$\dot{V} = L_K \times v_w. \tag{200}$$

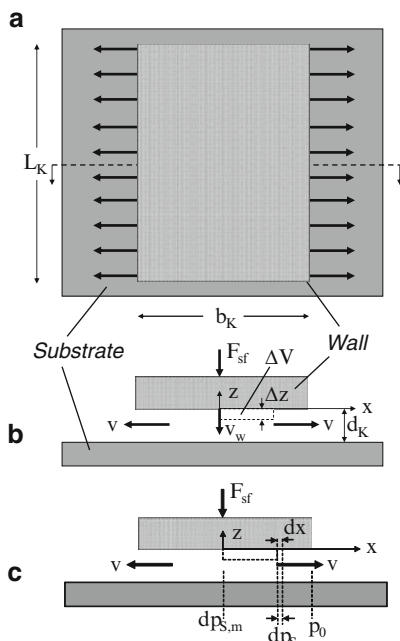
The flow velocity is obtained from this equation by dividing the cross-section of the capillary in the direction of the flow:

$$v = \frac{x}{d_K} v_w. \tag{201}$$

This velocity equals the velocity due to friction in a gap with length  $dx$  as a function of the driving pressure difference  $dp_S$  [(193) without the term of the capillary force, cf. Fig. 87c]:

$$v = \frac{x}{d_K} v_w = -\frac{d_K^2}{8 \eta dx} dp_S \Rightarrow dp_S = -\frac{8 \eta v_w}{d_K^3} x dx. \tag{202}$$

Integrating on both sides of this equation yields:



**Fig. 87** Top view (a) and cross-section (b) of a rectangular wall forming a capillary with the substrate. (c) Pressures in the capillary

$$\int_{p_{S,m}}^{p_S} dp'_S = \int_0^x -\frac{8 \eta v_w}{d_K^3} x' dx' \Rightarrow p_S(x) - p_{S,m} = -\frac{8 \eta v_w}{d_K^3} \frac{1}{2} x^2. \quad (203)$$

At the end of the capillary (at  $x = b_K/2$ ), the pressure is the pressure of the environment  $p_0$ . From this boundary condition, it is obtained:

$$p_S(x) = \frac{4 \eta v_w}{d_K^3} \left( \frac{b_K^2}{4} - x^2 \right) + p_0. \quad (204)$$

The force  $F_{sf}$  necessary to press the wall down with the velocity  $v_w$  can be calculated now from the integral over the force contributions  $dF_{sf}$  which are acting on the area with the length  $L_K$  and the width  $dx$ :

$$\begin{aligned} F_{sf} &= \int_{-b_K/2}^{b_K/2} dF_{sf} = \int_{-b_K/2}^{b_K/2} (p_S(x) - p_0) L_K dx \\ &= \int_{-b_K/2}^{b_K/2} \frac{4 \eta v_w}{d_K^3} \left( \frac{b_K^2}{4} - x^2 \right) L_K dx = \frac{2}{3} \frac{\eta L_K b_K^3}{d_K^3} v_w. \end{aligned} \quad (205)$$

From this force, the pressure can be calculated which is necessary to press the wall down:

$$\Delta p_{sf} = \frac{F_{sf}}{L_K b_K} = \frac{2}{3} \frac{\eta b_K^2}{d_K^3} v_w. \quad (206)$$

In a similar way, this calculation can be performed for the squeeze film effect generated by a circular wall with the radius  $R_W$ :

$$F_{sf} = \eta \pi \frac{R_W^4}{d_K^3} v_w \quad (207)$$

$$\Delta p_{sf} = \eta \frac{R_W^2}{d_K^3} v_w. \quad (208)$$

The squeeze film effect is important in microtechnique, because many microstructures are moving near the surface of a substrate or next to another structure. For example, the electrodes of the acceleration sensor shown in the exercises in Fig. E13 in problem 12 on page 82 are subdivided into small parts allowing air escaping between them. Thus, this acceleration sensor does not need to be evacuated to avoid damping by the squeeze film effect.

If a microstructure is moving parallel to a wall, there is also some friction, because the viscosity of the fluid hampers the flow in the gap which is caused by the moving body. The fluid in contact to the surface of the wall and the body nearly adheres to it. As a consequence, a linear flow profile is developing as shown in

Fig. 88. The friction force  $F_R$  of a flat body with area  $A_K$  and distance  $d_K$  to the wall is given by:

$$F_R = \eta \frac{A_K}{d_K} v_w. \quad (209)$$

Even if there is no wall in the near of the moving body, there is some friction due to the viscosity of the surrounding fluid, the so-called *Stokes friction*. If the body is a sphere with radius  $R_K$ , it can be calculated exactly with the following equation:

$$F_R = 6 \pi \eta R_K v_w. \quad (210)$$

The liquid in a capillary can be moved by *electroosmosis* also. If a polar liquid such as water is in contact with a rigid material, electrical charges are exchanged between liquid and rigid body. As a result, the rigid body is charged negatively and positive charges in the liquid are attracted to its surface (Fig. 89). A positively charged layer is formed which adheres to the surface. This layer is called the Stern layer. Adjacent to the Stern layer, a diffusion layer is built up which is caused both by the attraction of the negative charges and the diffusion of the constituents of the liquid due to the Brownian motion driven by temperature. The diffusion layer contains an excess of positive charges also. Stern layer and diffusion layer are called the electrical double layer (EDL).

An electrical potential  $\Phi$  is generated by the double layer which drops exponentially from the wall into the liquid:

$$\Phi = \Phi_W e^{-z/\lambda_D}. \quad (211)$$

In this equation,  $\Phi_W$  is the potential at the surface of the rigid body and  $\lambda_D$  is the so-called Debye–Hückel length which is a measure of the potential drop as a function of the distance  $z$  from the surface and the thickness of the diffusion layer:

$$\lambda_D = \sqrt{\frac{\epsilon_0 \epsilon_r k_B T}{2 q_e^2 e^2 c_F}}, \quad (212)$$

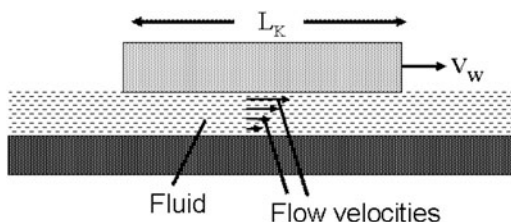
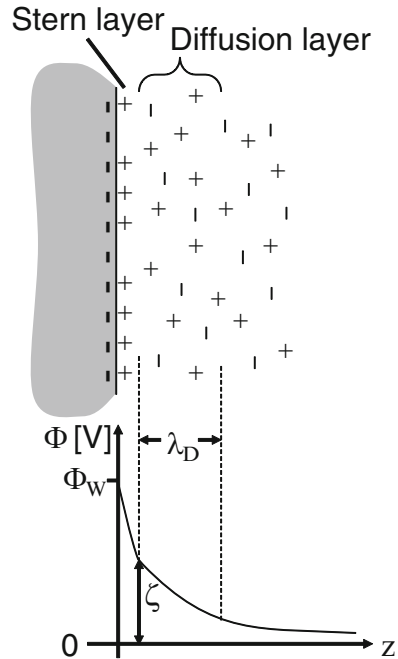


Fig. 88 Flow profile generated between a moving microstructure and a wall

**Fig. 89** Electrical double layer consisting of Stern and diffusion layer and electrical potential due to these layers



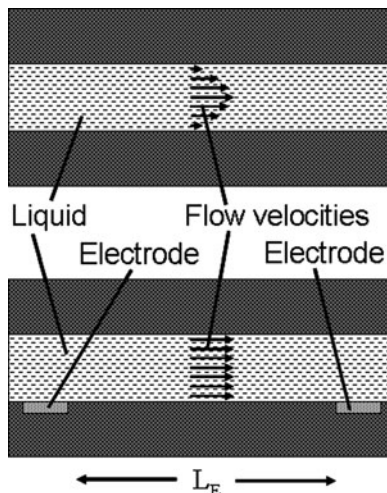
where  $\epsilon_0 = 8.9 \times 10^{-12}$  A s/(V m) and  $\epsilon_r$  are the absolute and relative permittivity, respectively,  $q_e$  is the number of elementary charges of the ions in the liquid (e.g.,  $q_e = -1$  for  $\text{OH}^-$  and  $q_e = 2$  for  $\text{Ca}^{2+}$ ),  $e = 1.6 \times 10^{-19}$  C is the elementary charge,  $c_F$  is the concentration of charges in the liquid at a position far from the surface of the rigid body (number of particles per  $\text{m}^3 = \text{mol}/\text{m}^3 \times N_A$ ),  $N_A = 6.02 \times 10^{23}/\text{mol}$  is the Avogadro constant,  $k_B = 1.4 \times 10^{-23}$  J/K is the Boltzmann constant, and  $T$  is the absolute temperature measured in Kelvin (K).

The electrical potential  $\zeta$  at the interface between Stern layer and diffusion layer is called the  $\zeta$ -potential. The Stern layer adheres to the surface of the rigid body, but the diffusion layer can be moved by an electric field. Therefore, electrodes in a capillary move the diffusion layer, and due to friction in the liquid, the charged particles in the diffusion layer take the neutral particles along.

The friction is largest at the wall of the capillary and the electroosmotic force is largest at the wall also. Therefore, these two effects approximately balance one another, and, as a consequence, the flow velocity is approximately constant over the channel width, if it is not too wide (cf. Fig. 90). This is different from the parabolic flow profile of a flow driven by a pressure difference over a capillary, and this is an important condition for analyzing different ingredients by electrophoresis (see below).

If the EDL is thin compared with the width of the capillary and the electric field is not too large, the flow velocity  $v$  of a liquid with dynamic viscosity  $\eta$  and  $\zeta$ -potential

**Fig. 90** Flow profile generated by a pressure drop (*top*) and by electroosmosis (*bottom*)



$\zeta$  generated by electroosmosis between electrodes with a distance  $L_E$  and charged with a voltage  $U$  can be calculated with the Helmholtz–Schmoluchovski equation:

$$v = \frac{\varepsilon_0 \varepsilon_r E_{el} \zeta}{\eta} = \frac{\varepsilon_0 \varepsilon_r U \zeta}{\eta L_E}, \quad (213)$$

where  $E_{el}$  is the electrical field between the electrodes. This flow velocity is generated by an electroosmotic pressure  $\Delta p_{EOF}$  which is equal to the pressure difference  $\Delta p_R$  necessary to overcome the friction in the capillary [(175) on page 107]:

$$\Delta p_{EOF} = \Delta p_R = 32 \frac{\eta L_E}{D_h^2} v = \frac{32}{D_h^2} \varepsilon_0 \varepsilon_r U \zeta. \quad (214)$$

The electroosmotic flow calculated with (213) is a function of the electrochemical interaction between liquid and rigid body, temperature, and concentration. As a consequence, the prediction of the velocity is not very reliable, because, e.g., a very thin layer of any molecules or atoms on the surface of the wall can change the  $\zeta$ -potential or a comparatively small change of the temperature will change both the  $\zeta$ -potential and the viscosity.

The volume flow  $\Phi_P$  generated by electroosmosis is the product of the cross-section  $A$  of the microchannel and the flow mean velocity  $v$  calculated with (213). In general, a counter pressure may reduce the electroosmotic flow. Thus, the entire pressure drop  $\Delta p_P$  over the flow channel is the difference of the electroosmotic pressure  $\Delta p_{EOF}$  and the pressure difference  $\Delta p_R$  necessary to overcome the friction in the capillary [(175) on page 107]:

$$\Delta p_P = \Delta p_{\text{EOF}} - \Delta p_R = \frac{32}{D_h^2} \varepsilon_0 \varepsilon_r U \zeta - 32 \frac{\eta L_E}{D_h^2} v = \frac{32}{D_h^2} \left( \varepsilon_0 \varepsilon_r U \zeta - \eta L_E \frac{\Phi_P}{A} \right) \quad (215)$$

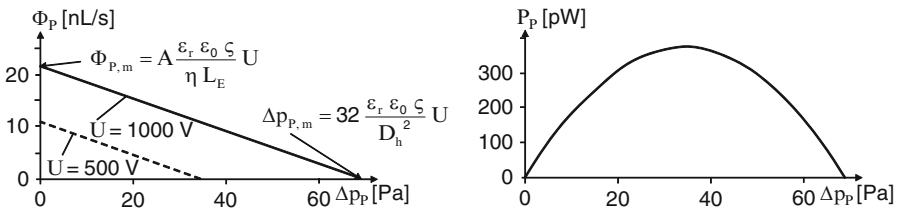
$$\Rightarrow \Phi_P = \frac{A}{\eta L_E} \left( \varepsilon_0 \varepsilon_r U \zeta - \frac{D_h^2}{32} \Delta p_P \right). \quad (216)$$

A channel between the electrodes in which an electroosmotic flow  $\Phi_P$  is generated is a kind of micropump. The characteristic curve of this pump is the volume flow generated as a function of the counter pressure  $\Delta p_P$  against which the pump is working. On the left of Fig. 91, the characteristic curve of an electroosmotic micropump with a channel length of 1 cm, a cross-section of  $100 \mu\text{m} \times 100 \mu\text{m}$ ,  $\zeta$ -potential of 30 mV, viscosity of 1 mPa s, and relative permittivity of 81 is shown as calculated with (216). The characteristic curve is similar to those of other micropumps driven by, e.g., a piezo (cf. Fig. 182 on page 238). However, both the flow and the pressure generated by an electroosmotic micropump are comparatively small. The maximum flow  $\Phi_{P,m}$  which can be achieved is limited by the maximum electric field which can be applied before the breakdown voltage is reached. A larger maximum counter pressure  $\Delta p_{P,m}$  can be achieved at a certain electrical field (the ratio of voltage  $U$  and distance of the electrodes  $L_E$ ) when the channel is longer. A compromise is necessary between the maximum flow and the maximum counter pressure, when the cross-section  $A$  of the channel is designed. A larger cross-section results in a larger maximum flow and also in a smaller maximum counter pressure, because the hydraulic diameter  $D_h$  increases when the cross-section is enhanced.

The power  $P_P$  generated by a pump is the product of the generated flow  $\Phi_P$  and the counter pressure  $\Delta p_P$ . The power of the electroosmotic pump with the characteristic curve shown on the left of Fig. 91 is shown on its right. The power generated by an electroosmotic pump generally is rather small (cf. Fig. 183 on page 239).

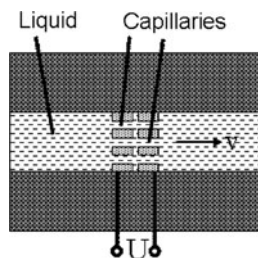
The electrical resistance of the electroosmotic pump is given by (432) on page 257.

The electroosmotic effect can also be reversed. A flow through an arrangement of parallel capillaries generates a voltage which is a function of the flow velocity as shown in Fig. 92. This voltage is produced by the charges transported together with



**Fig. 91** Characteristic curve of a electroosmotic micropump (left) and power  $P_P$  generated by such a pump driven with 1,000 V

**Fig. 92** Voltage generated by the reversed electroosmotic effect



a part of the diffusion layer which is taken along with the flow. This effect could even be employed as a flow sensor.

While the electroosmotic effect is moving the entire liquid in a capillary, *electrophoresis* is generating an additional flow velocity of single particles inside of the liquid. An EDL is also developing around particles in a liquid. The diffusion layer around the particles contains an excess of positive charges and is attracted to the cathode taking along the particle. The additional velocity  $v$ , which a particle gains this way, is calculated with the Hückel equation, if it is much smaller than the Debye–Hückel length  $\lambda_D$ :

$$v = \frac{2}{3} \frac{\varepsilon_0 \varepsilon_r E_{el} \zeta}{\eta} = \frac{2}{3} \frac{\varepsilon_0 \varepsilon_r U \zeta}{\eta L_E}. \quad (217)$$

Examples of such small particles are organic molecules and nanoparticles.

If the particles are larger than the Debye–Hückel length, the  $2/3$  in above equation needs to be omitted. Examples are cells or polymer spheres with a diameter of several micrometers, so-called beads.

If particles carry a charge  $q_P$  themselves such as ions, they are moved by the electric field also. An equilibrium is reached between the electrical force  $F_{el}$

$$F_{el} = q_P E_{el} = q_P \frac{U}{L_E}, \quad (218)$$

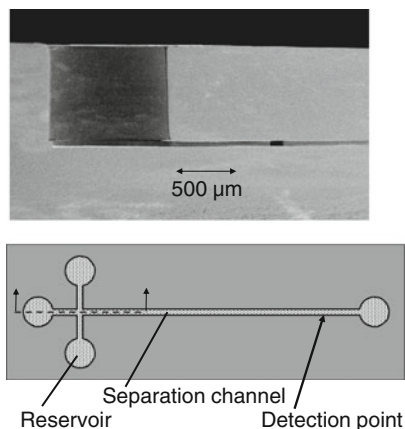
and the friction force  $F_R$  of a sphere moving in a fluid, the Stokes equation (210):

$$F_R = 6 \pi \eta R_T v \Rightarrow v = \frac{q_P U}{6 \pi \eta R_T L_E}. \quad (219)$$

Here  $R_T$  is the radius of the particle and the surrounding hydrate hull which is moved together with it.

Different kinds of particles in a liquid show different  $\zeta$ -potentials, charges, and radii. Therefore, they move with different velocities and are separated from each other, if they start at the same time from a common position in a capillary. This effect is employed for chemical and biological analyses in microchannels. Figure 93

**Fig. 93** Microchannel for capillary electrophoresis. *Top:* Cut through a part of a channel (Courtesy of Karlsruhe Institute of Technology, KIT); *Bottom:* Schematic top view



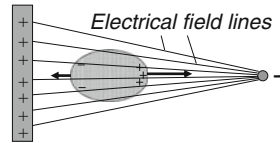
shows the principal design of such a device. An electrophoresis chip contains a cross-like capillary system filled with a gel. At the four ends of the channels, reservoirs with access to the environment are placed. The sample to be analyzed is introduced at the reservoir of one of the short arms of the capillary system. Electrodes are immersed into the reservoirs and a voltage on the order of several hundred volts between the ends of the two short arms is transporting the sample across the intersection point of the channels by electroosmosis and electrophoresis. Then a voltage is applied over the long separation channel of the chip and the ingredients of the part of the sample at the intersection point are moved along the separation channel. The difference in velocity of the electrophoretic flow of ingredients of different type separates them from each other and they arrive at a detection point down stream at different times.

Ingredients are distinguished by the time they need to arrive at the detection point and their amount is calculated from the size of the peak developing when it is passing the detection point. Detection needs to be sensitive to almost every possible ingredient. UV transmission, fluorescence, and conductivity of the liquid are used mostly.

Particles inside of a liquid can be moved by alternating voltage, also, if the electrical field is asymmetric. This phenomenon is called *dielectrophoresis*. An explanation of this phenomenon is that charges are induced on the particle by the electrical field. More charges are induced where the field is stronger, and the particle is attracted towards the gradient of the electric field (cf. Fig. 94). If the polarity of the electrical field is changed, the resulting forces are the same. Dielectrophoretic forces may be repulsive also. This can be explained by a buoyancy force in the electrical field which is generated, because the liquid is attracted more than the particle. It remains to be difficult to understand that an arrangement of electrodes may change between attracting and repulsing particles just by changing the frequency. The voltages applied for dielectrophoresis are on the order of some 10 V with a frequency in the range of MHz.



**Fig. 94** Particle moved by dielectrophoresis by the gradient of the electric field



**Table 12** Equations for the calculation of pressures acting in capillaries

Type	Equation
Pressure difference overcoming friction	$\Delta p_R = -32 \frac{\eta L_F}{D_h^2} v$
Pressure difference accelerating to velocity $v$	$\Delta p = -\frac{\rho_F}{2} v^2$
Capillary pressure	$\Delta p_K = -\frac{U_w}{A_F} \Gamma \cos(\theta_w)$
Pressure overcoming squeeze-film effect at a rectangular wall (rough approximation)	$\Delta p_{sf} = \frac{2}{3} \frac{\eta b_K^2}{d_K^3} v_w$
Pressure overcoming squeeze-film effect at a circular wall (rough approximation)	$\Delta p_{sf} = \eta \frac{R_W^2}{d_K^3} v_w$
Electroosmotic pressure	$\Delta p_{EOF} = \frac{32}{D_h^2} \varepsilon_0 \varepsilon_r U \zeta$
Capacitive pressure on dielectric liquid	$\Delta p_C = \frac{1}{2} \frac{\varepsilon_0}{d_i^2} \Delta \varepsilon_r U^2$
Capacitive pressure on conductive liquid	$\Delta p_C = \frac{1}{2} \frac{\varepsilon_0}{d_i^2} \Delta \varepsilon_r U^2$

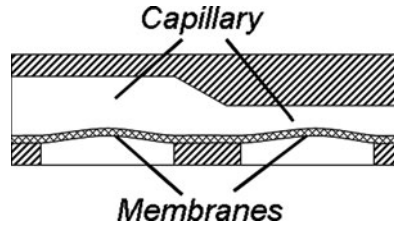
A liquid can be attracted into or ejected from a partly filled capillary by capacitive forces also. This is described in the following chapter, but the corresponding equations are included in Table 12, which summarizes all equations describing forces which move fluids in a capillary.

## Exercises

### Problem 17

In Fig. E18, the working principle of a flow sensor is shown. Strain gauges are mounted on the two membranes at the circumference of the capillary. A bridge circuit measures the difference of the deflection of the two membranes only. Temperature changes and pressure changes outside of the sensor which affect the deflection of both membranes and the electrical resistance of all resistances are compensated by the bridge circuit.

**Fig. E18** Capillary with a changing height and two membranes



- (a) Derive an equation with which the flow velocity (in m/s) can be calculated from the measured pressure difference.
- (b) Use the equation from (a) to calculate the pressure difference which is the result of a water flow of 1 m/s in the wider part of the capillary, if the height of the capillary at the constriction is reduced by a factor of 2.
- (c) If the membrane at the right is placed behind the constriction where the capillary gets as wide as at the position of the other membrane, what pressure difference is measured in that case?

Dynamic viscosity of water	1.002 mPa s	Density of water	1,000 kg/m <sup>3</sup>
Length of the constriction	1 mm	Height and width of the constriction	200 μm

**Problem 18**

The rising of a liquid in a capillary is a function of the surface tension and the wetting angle.

- (a) Two glass plates are bonded to each other in such a way that there is a gap between them. These plates are immersed into a cup filled with water. Please, calculate the height up to which the water rises in the gap of 0.5 mm, 20 μm, and 2 μm at 20°C, respectively. The width of the gap is 50 mm. Is the rising height of the water a function of the gap width?  
 Hint: Consider the equilibrium of forces between the weight and the capillary force. Express the weight in the capillary gap as:  
 $F_G = \rho_w b_K d_K L_K g_e$ , where  $\rho_w$  is the density of water ( $\rho_w = 1 \times 10^3 \text{ kg/m}^3$ ),  $b_K$  is the width of the gap, and  $g_e = 9.81 \text{ m/s}^2$  is the acceleration of gravity.
- (b) Calculate the flow velocity in the unit [mm/s] in the three gaps given above, if the gaps are lying horizontally and are filled to 1 cm.
- (c) Have you taken into account for your calculation in (b) that the water needs to be accelerated to achieve its velocity? Please calculate the mean flow velocity with and without the effect of the acceleration.
- (d) What is the rising height for the three cases of (a) at a water temperature of 80°C?
- (e) How does the rising height change if a surface-active agent is mixed into the water which reduces the surface tension to 5 mN/m? The wetting angle changes to 10°. Please calculate the rising height for the three gap sizes given in (a).

- (f) How do the rising heights change, if the plates are not made of glass but from silicone? For example, a structure could be cast of PDMS (poly dimethyl siloxan). PDMS is a two-component silicone which is used in microtechnique to mold surface structures.

Please calculate the rising height for silicone plates and the three gap sizes in (a).

Surface tension of pure water at 20°C	72.7 mN/m	Surface tension of pure water at 80°C	62.5 mN/m
Wetting angle of pure water on glass	14°	Wetting angle of pure water on silicone	110°
Dynamic viscosity of pure water	1.002 mPa s		

### ***Problem 19***

A capillary is 5 cm long, 100 μm high, and at both ends there are electrodes. It can be filled with a water solution without bubbles.

- (a) Please calculate the flow velocity inside of the capillary, if a voltage of 1,000 V is supplied to the electrodes.
- (b) Please calculate the additional velocity with which the amino acid glycine is transported in the capillary.
- (c) Please calculate the electrophoretic velocity of sodium ions in the same setup. Use the equation of the velocity which was derived from the Stokes equation (210 on page 122). Assume 0.2 nm for the radius of the hydrate hull of sodium.
- (d) Electrophoresis is employed in bioscience to separate molecules. In certain analytic applications at the end of a capillary, there is a detector arranged detecting the molecules (e.g., by fluorescence).

How long needs a separation channel to be which shall be suitable for the separation and detection of glycine and sodium? The time sensitivity of the detector is assumed to be 0.1 s. How much time is required for separation?

Absolute permittivity	$8.9 \times 10^{-12}$ A s/(V m)	Relative permittivity of water	81
ζ-potential of water	30 mV	ζ-potential of glycine	70 mV
Dynamic viscosity of the water solution	1.002 mPa s		

# Capacitive Forces

Capacitive forces transform a voltage directly into a movement and they are insensitive to temperature changes. These are the main reasons why capacitive forces are employed so often in microtechnique.

A capacitor consists of two electrodes mounted at a certain distance. When the electrodes are charged, they are attracted to each other due to the electrostatic Coulomb force. There are several kinds of capacitive forces distinguished by the way of movement which is generated by them. They all can be calculated from the potential energy  $W_C$  stored in the capacitor which is a function of the voltage  $U$  and the capacity  $C_{el}$ :

$$W_C = \frac{1}{2} C_{el} U^2. \quad (220)$$

Here,  $C_{el}$  is the electrical capacity which is a function of the absolute  $\epsilon_0$  and the relative  $\epsilon_r$  permittivity of the material between the electrodes, their area  $A_C$ , and the distance  $d_C$  between the electrodes:

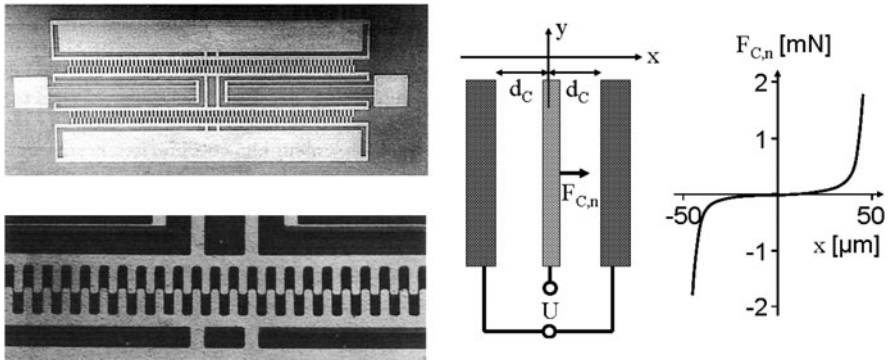
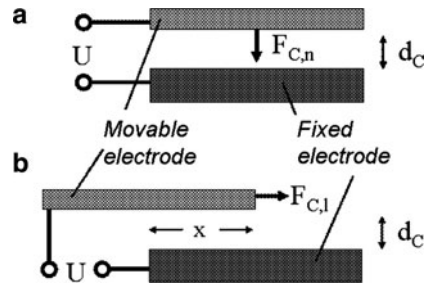
$$C_{el} = \epsilon_0 \epsilon_r \frac{A_C}{d_C} \Rightarrow W_C = \frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C} U^2. \quad (221)$$

In general, the force is the derivative of the potential energy with respect to the path of the movement. This method has been introduced by Ritz to calculate the deformation of bodies (cf. page 34f). Thus, if the bearing of the electrodes is arranged such that one electrode can move towards the other, a capacitive force  $F_{C,n}$  is acting normal (perpendicular) to the electrodes (cf. Fig. 95a) and it is calculated as the derivative of the energy with respect to the distance of the plates:

$$F_{C,n} = \frac{\partial W_C}{\partial d_C} = -\frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C^2} U^2. \quad (222)$$

The normal capacitive force is inversely proportional to the square of the distance of the electrodes. Therefore, large forces can be obtained only if the electrodes are near to each other limiting the possible stroke of this kind of actuator.

**Fig. 95** Normal (a) and lateral (b) capacitive forces



**Fig. 96** *Left:* Comb drive [5] Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission. *Right:* Movable electrode between two fixed ones and capacitive force acting on it as calculated with (225)

If the bearing of the movable electrode allows a movement in lateral direction only, the capacitive force  $F_{C,l}$  needs to be calculated from the derivative with respect to the overlapping length of the electrodes and the area is expressed as the product of overlapping length  $x$  and width  $b_C$  of the electrodes (cf. Fig. 95b):

$$F_{C,l} = \frac{\partial W_C}{\partial x} = \frac{1}{2} \varepsilon_0 \varepsilon_r \frac{b_C}{d_C} U^2. \quad (223)$$

This lateral capacitive force is not a function of the position of the electrode. Thus, the force is constant until it vanishes when the electrodes are overlapping completely. This capacitive force is employed for comb drives which allow for a larger actuator stroke with a constant force enlarged by multiple electrodes. Figure 96 shows on the left such a comb drive.

At comb drives, the normal capacitive force is acting also. If the bearing [typically beams clamped at one end or folded beams (cf. page 77)] is not strong enough, the electrodes may touch each other resulting in a short circuit. The potential energy of an electrode between two other electrodes is the sum of the energies stored in the capacitors on both sides of the movable electrode (cf. Fig. 96). From (221) it is derived:

$$W_C = \frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C + x} U^2 + \frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C - x} U^2 = \frac{1}{2} \epsilon_0 \epsilon_r A_C U^2 \left( \frac{1}{d_C - x} + \frac{1}{d_C + x} \right). \quad (224)$$

The normal force acting on the movable electrode again is calculated by the derivative of the energy:

$$F_{C,n} = \frac{\partial W_C}{\partial x} = \frac{1}{2} \epsilon_0 \epsilon_r A_C U^2 \left( \frac{1}{(d_C - x)^2} - \frac{1}{(d_C + x)^2} \right). \quad (225)$$

Figure 96 shows on the right the normal capacitive force of a movable electrode calculated with (225). The area of the electrode is  $A_C = 1 \text{ mm}^2$ , its distance to the fixed electrodes is in the center position  $d_C = 50 \text{ }\mu\text{m}$ , and it is powered with 100 V. In the middle between the electrodes, there is an unstable equilibrium of forces, because the movable electrode is attracted to both sides with the same force. If the position is changed towards one side, the force is increasing strongly. Therefore, in the near of the center position not much force is needed to balance the electrode of a comb drive but stoppers or some other means is needed avoiding that the central position is left too much, e.g., by an impact from the outside.

The lateral capacitive force as calculated with (223) can be increased if the movable electrode is not a rigid body but an electrically conductive liquid as schematically shown in Fig. 97a. Since one of the electrodes is the conductive liquid now, the distance of the electrodes is reduced in this case to the thickness  $d_l$  of an insulating layer on the electrode and the relative permittivity is the one of that layer. For this calculation, it is neglected that the part of the capacitor not filled with liquid is also contributing a bit to the total stored energy, reducing the force a little bit.

Even if the liquid or a rigid body between two electrodes is insulating (Insulating material is called dielectric.), there is a capacitive force attracting it into the capacitor.

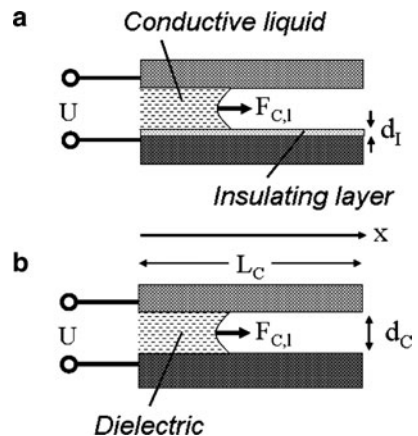


Fig. 97 Lateral capacitive force on (a) a conductive liquid and (b) a dielectric

This is due to atoms or molecules in the material acting as electrical dipoles which align to the electrical field in the capacitor and are attracted this way. This effect is a function of the relative permittivity of the material. The capacitor can be considered consisting of two parts filled with different dielectrics. Therefore, the potential energy stored in the capacitor is calculated as the sum of the energies of both parts:

$$W_C = \frac{1}{2} \varepsilon_0 \varepsilon_{r,1} \frac{b_C x}{d_C} U^2 + \frac{1}{2} \varepsilon_0 \varepsilon_{r,2} \frac{b_C (L_C - x)}{d_C} U^2. \quad (226)$$

Calculating the derivative with respect to  $x$  yields the force:

$$F_{C,1} = \frac{\partial W_C}{\partial x} = \frac{1}{2} \varepsilon_0 \frac{b_C}{d_C} U^2 (\varepsilon_{r,1} - \varepsilon_{r,2}). \quad (227)$$

Note that this capacitive force may also become repulsive (negative) when the relative permittivity of the second material is larger than that of the first one. In such a case, the material with the smaller permittivity is ejected from the capacitor.

The pressure  $\Delta p_C$  generated in a capillary when a dielectric liquid is attracted into it can be calculated from (227) by dividing by the cross-sectional area:

$$\Delta p_C = \frac{1}{2} \frac{\varepsilon_0}{d_C^2} U^2 (\varepsilon_{r,1} - \varepsilon_{r,2}). \quad (228)$$

The capacitive forces are compared with each other in Fig. 98. For all calculations, electrodes with a length of  $L_C = 2$  mm, a width of  $b_C = 0.5$  mm, and a voltage supply of 100 V were assumed. The normal and lateral forces were calculated with (222) and (223), respectively, assuming air between the electrodes ( $\varepsilon_r = 1$ ). For the lateral forces, a distance of 100  $\mu\text{m}$  between the electrodes was used. For the conductive liquid, the thickness and relative permittivity of the insulation layer were set to 1  $\mu\text{m}$  and 2.5, respectively.

Figure 98 shows that for small distances between the electrodes the normal force is much larger than all other capacitive forces. Since the normal force is

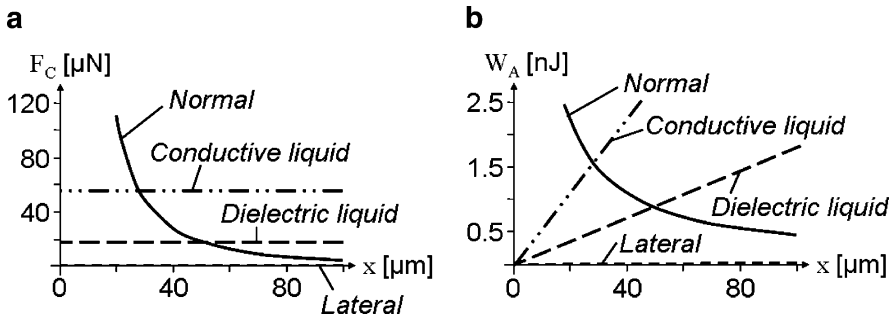


Fig. 98 (a) Capacitive forces  $F_C$  and (b) work  $W_A$  as a function of the stroke  $x$

proportional to the inverse of the square of the distance, at larger distances the lateral forces are dominating. Especially, the conductive liquid in a capacitor allows for comparatively large forces, because the thickness of the insulating layer may be manufactured very thin compared with the required height of a capillary.

The force acting on a dielectric liquid may be comparatively large, if a proper dielectric has been chosen. For the calculation shown in Fig. 98, water ( $\epsilon_r = 81$ ) was assumed to be attracted into a channel filled with air.

The lateral capacitive force on a single electrode is comparatively small as shown in Fig. 98. This is the reason why comb drives need such a large number of electrodes. There are two unique advantages of lateral capacitive actuators. One of them is that the force is not a function of the stroke. Therefore, the stroke can be designed very large. This is illustrated by the work  $W_A$  done by the actuator as a function of the stroke length which is shown on the right of Fig. 98. The work is the integral of the force over the stroke length. Since the lateral force is constant with respect to the stroke, the work done by the actuator rises linearly with the stroke.

The other unique advantage of capacitive actuators is that they show nearly no cross-sensitivity to temperature changes. The characteristic curves of almost all sensors and actuators are a function of temperature. Even the elastic deflection of a beam [cf. (97) on page 66] varies with temperature because Young's modulus is a function of temperature. So, cross-sensitivity to temperature changes is one of the major problems (not only) in microtechnique. Whenever a design is made, the possible effect of temperature changes needs to be considered. The behavior of capacitive actuators changes only due to thermal strain which is generally smaller than other temperature effects such as the change of piezoelectric coefficients (cf. Fig. 114 on page 153) or the influence on thermal actuators for which the deflection is a direct effect of the temperature change.

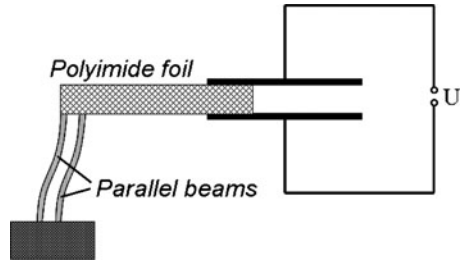
The efficiency of electrostatic actuators is very large and close to 100%. This is an advantage especially when battery powered devices need to be designed with small energy consumption. Besides the efficiency, the maximum energy output density is an interesting quantity, because microactuators typically are not able to provide large energy output. The maximum energy output density is the total power output which can be generated with an actuator per volume when there are no restrictions on the energy consumption. Since the efficiency of electrostatic actuators is very large, the energy output can be assumed to be approximately the same as the energy stored in the capacitor which is calculated with (221). The volume  $V$  of the capacitor is approximately its area  $A_C$  times the sum of the distance  $d_C$  of the electrodes and their thickness  $d_E$ . Therefore, the energy output density is:

$$\frac{W_A}{V} = \frac{1}{2} \epsilon_0 \epsilon_r \frac{U^2}{d_C d_E}. \quad (229)$$

If a maximum voltage of 400 V and a distance of the electrodes of 10  $\mu\text{m}$  are assumed as possible maximum and minimum values, respectively, and the needed height of the actuator is 100  $\mu\text{m}$ , the maximum energy density of an electrostatic actuator is 0.7  $\mu\text{J}/\mu\text{L}$ .



**Fig. E19** Capillary with laterally movable polyimide foil fixed at parallel beams



## Exercises

### Problem 20

A thick foil from polyimide can move in the space between the electrodes of a capacitor (cf. Fig. E19). The foil is fixed at two parallel beams.

The distance between the capacitor electrodes is  $20\ \mu\text{m}$ . The width of the capacitor plates is  $0.5\ \text{mm}$ . The thickness of the polyimide foil is  $20\ \mu\text{m}$ . The beams are  $5\ \text{mm}$  long,  $10\ \mu\text{m}$  wide, and  $10\ \mu\text{m}$  thick. They are made of silicon with a Young's modulus of  $120\ \text{GPa}$ .

- Please calculate the force acting on the polyimide foil, when a voltage of  $50\ \text{V}$  is applied to the electrodes. In which direction is the force acting on the foil? This microsystem is surrounded by air.  
Hint: Assume that the thickness of the foil is the same as the distance of the plates, but there is no friction between foil and electrodes.
- Derive an equation for the calculation of the position of the polyimide foil as a function of the voltage applied. What is the position of the foil when  $50\ \text{V}$  are applied?
- The microsystem is now placed into a cup filled with pure water. In which direction does the force act on the foil now, when a voltage is applied? Calculate this force.
- What is the force if the lower electrode is fixed to the beams and neither polyimide nor water is in the capacitor?

Absolute permittivity	$8.9 \times 10^{-12}\ \text{A s}/(\text{V m})$	Relative permittivity of polyimide	3.4
Relative permittivity of air	1.0	Relative permittivity of water	81

# Piezoelectric Effect

The piezoelectric effect is widely employed in microtechnique. Especially, piezos used as an actuator are appreciated because of the large force which can be generated in a small volume. But piezos can be employed also as sensors which provide a large output voltage.

If a piezoelectric material is loaded with a pressure, it is compressed as any other elastic material, and it is extended transversal to the direction of the load (cf. Fig. 99b). In addition, electrical charges are generated on electrodes on the surface of the piezoelectric material and it is possible to measure a voltage between the electrodes. The generation of charges is called the piezoelectric effect and the material is called a piezo. The origin of the word “piezo” is Greek and it means “I am pressing”. The piezoelectric effect is employed to measure strains, forces, or pressures.

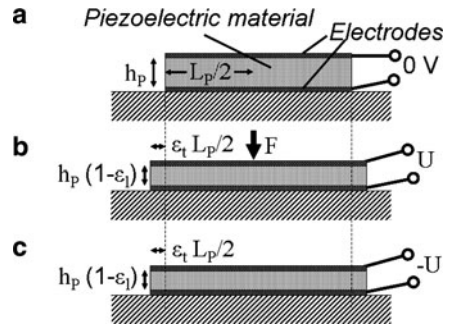
There is also an inverse piezoelectric effect: When a voltage is supplied to the electrodes, the piezo is straining both longitudinal and transversal to the electric field as shown in Fig. 99c. This inverse piezoelectric effect is employed to build actuators.

There are two kinds of piezoelectric materials: *monocrystalline* materials, such as quartz, zink oxide, and lithium niobate, and *ferroelectric* materials, such as the ceramics PZT (lead–zirconate–titanate) and barium titanate, and the polymer PVDF (poly vinyliden fluoride). The piezoelectric effect of monocrystalline materials is a function of crystal orientation. In general, it is comparatively small. Therefore, the piezoelectric effect of ferroelectric materials is employed in most cases.

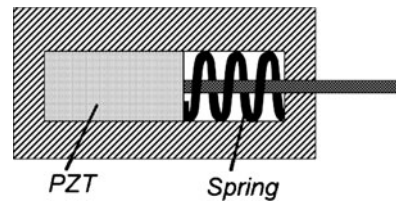
Ferroelectric materials are not yet piezoelectric after they have been fabricated. They need to be polarized by a large electrical field before use; similar to ferromagnetic materials which need to be magnetized in a magnetic field before they become permanent magnets. This analogy is the reason for the name “ferro”electric, although these materials have nothing in common with iron.

Also in analogy to ferromagnetic materials, ferroelectric materials have a Curie temperature. If they are heated up over this temperature, the piezoelectric effect disappears again. Again in analogy to ferromagnetic materials, ferroelectric materials can be depolarized by an electric field opposite to their direction of polarization. As a consequence, ferroelectric materials should not be exposed to high temperatures and large electric fields opposite to their polarization.

**Fig. 99** A piezo pressed together generates electrical charges on its electrodes (b), and a voltage supplied to the electrodes is straining the piezo. In (a) the piezo is shown without mechanical load and discharged



**Fig. 100** Piezo mounted such that it is always under compressive stress



It needs to be noted also that the most commonly used piezoelectric material PZT, which shows the largest effect, is a ceramic. As a consequence, it is brittle and tends to break when loaded with a tensile force. Tensile stress in a piezo can be avoided by designing an actuator such that it is under compressive stress always. Figure 100 shows an example.

Usually, the directions of a piezo are denominated by the subscripts 1, 2, and 3. The direction of the polarization of the piezo is marked with the subscript 3. Typically, the direction of polarization is identical with the connecting line of the electrodes, because the piezo was polarized by applying a large voltage to these electrodes.

The strain  $\varepsilon_3$  in polarization direction of a piezo is calculated with the following equation:

$$\varepsilon_3 = d_{33} E_{el} + \frac{\sigma_3}{E_P}, \quad (230)$$

where  $E_{el}$  is the electrical field and  $\sigma_3$  is the stress acting in polarization direction.  $E_P$  is Young's modulus of the piezo when the electrodes are short-circuited. The quantity  $d_{33}$  is the piezoelectric modulus of the piezo.  $d_{33}$  is a material property, which denotes the strength of the piezoelectric effect.

Equation (230) consists of two terms. The second one is already known; it is the strain of an elastic material according to Hooke's law. The first term in (230) represents the strain generated by the inverse piezoelectric effect.

The deflection  $w_3$  in polarization direction of a piezo is obtained from (230) by multiplying with the thickness  $h_p$  of the piezo which in this case is equal to the distance between its electrodes:

$$h_p \varepsilon_3 = w_3 = d_{33} h_p E_{el} + h_p \frac{\sigma_3}{E_p} = d_{33} U + \frac{h_p}{E_p A_p} F_3 = d_{33} U + F_3/k_3. \quad (231)$$

The last part of the equation was obtained with the definition of the homogeneous electrical field  $E_{el}$  in a capacitor as the quotient of the voltage and the distance of the electrodes, and the force  $F_3$  acting in polarization direction was calculated as the product of the homogeneous stress  $\sigma_3$  and the cross-sectional area  $A_p$  of the piezo perpendicular to the stress. The factor behind  $F_3$  is the spring constant  $k$  of an elastic body.

The unit of the piezoelectric modulus obviously is (m/V). The first index denotes the direction in which the voltage is applied and the second stands for the direction in which the deflection is calculated. Thus, the strain in 1-direction (the direction of the length  $L_p$  of the piezo) under the action of forces in 1- and 3-direction is calculated with:

$$\varepsilon_1 = d_{31} E_{el} + \frac{\sigma_1}{E_p} - \nu_p \frac{\sigma_3}{E_p}, \quad (232)$$

where  $\nu_p$  is Poisson's ratio of the piezo. The deflection in 1-direction is obtained by multiplying with the length  $L_p$  of the piezo:

$$L_p \varepsilon_1 = w_1 = d_{31} \frac{L_p}{h_p} U + \frac{L_p F_1}{E_p h_p b_p} - \nu_p \frac{F_3}{E_p b_p}, \quad (233)$$

where  $b_p$  is the width of the piezo. The absolute value of the piezoelectric modulus  $d_{31}$  is typically approximately half of the one of  $d_{33}$ . Therefore, it might be concluded that employing the effect in polarization direction would be preferable compared with the effect in transversal direction; but the opposite is true, because the length  $L_p$  of a piezo typically is much larger than its thickness  $h_p$ , and the factor  $L_p/h_p$  is much larger than two.

The electric field might also be applied in 1-direction transversal to the polarization and the deflection in 1-direction needs to be calculated. In this case, the piezoelectric modulus  $d_{11}$  needs to be employed:

$$\varepsilon_1 = d_{11} E_{el} + \frac{\sigma_1}{E_p} - \nu_p \frac{\sigma_3}{E_p} = d_{11} \frac{U}{L_p} + \frac{\sigma_1}{E_p} - \nu_p \frac{\sigma_3}{E_p} \quad (234)$$

$$\Rightarrow w_1 = L_p \varepsilon_1 = d_{11} U + L_p \left( \frac{\sigma_1}{E_p} - \nu_p \frac{\sigma_3}{E_p} \right). \quad (235)$$

However, applying a voltage perpendicular to the polarization is very difficult, because after fabrication and polarization of the piezo, electrodes need to be

manufactured in 1-direction which are not short-circuited with the electrodes employed for polarization. Therefore,  $d_{33}$  and  $d_{31}$  are sufficient to calculate the deflection of a piezo in most cases.

Equations (231) and (233) are employed for the calculation of the deflection of a piezo as an actuator. Figure 101 shows the deflection  $w_3$  as a function of the force  $F_3$ , calculated with (231) for a piezo from PZT with a Young's modulus and piezoelectric modulus of 60 GPa and  $250 \times 10^{-12}$  m/V, respectively, an area of  $10 \text{ mm}^2$ , and a distance  $h_P$  between the electrodes of 0.5 mm. This graph is the characteristic curve of a piezo as an actuator. The maximum force  $F_0$  generated by the piezo is achieved when it is hindered from deflecting, and no force can be generated at the maximum deflection  $w_E$  which can be generated. Between maximum force and maximum deflection, there is a linear curve which even extends over the points of maximum deflection and maximum force, because an external force may further deflect the piezo.

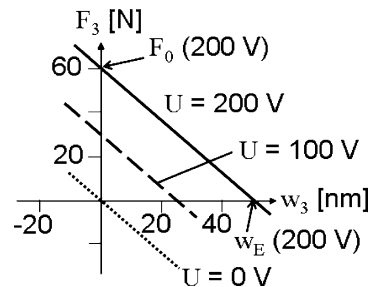
The maximum deflection  $w_E$  of a piezo, which is obtained when no external force is acting on it, is calculated by inserting no force in (231). The maximum force  $F_0$  which can be generated by a piezo is obtained when it is hindered from deflecting. In polarization direction, this is calculated by setting the deflection in (231) to zero and to solve for the force  $F_3$ :

$$F_0 = -d_{33} \frac{E_P A_P}{h_P} U. \quad (236)$$

If the voltage supplied to the piezo is changed, the characteristic curve is shifted as shown in Fig. 101. Without any voltage supplied, the characteristic curves simply represents the deflection due to Hooke's law (cf. 231).

Equations (231) and (233) are used to calculate the deflection of a piezo employed as an actuator. If a piezo is to be used as a sensor, the voltage generated by the deformation of the piezo needs to be calculated. The charge density  $D_3$  generated in 3-direction is obtained from the following equation:

$$D_3 = d_{33} \sigma_3 + \epsilon_0 \epsilon_r E_{el}. \quad (237)$$



**Fig. 101** Characteristic curves of a piezo as an actuator driven at various voltages

The charge  $Q_3$  on the electrodes in polarization direction of the piezo is obtained from this equation by multiplying with the area  $A_P$  of the electrodes ( $A_P$  is equal to the area of the piezo in this case):

$$Q_3 = d_{33} F_3 + \frac{\varepsilon_0 \varepsilon_r A_P}{h_P} U = d_{33} F_3 + C_{el} U. \quad (238)$$

The above equation consists of two terms. The second one represents the charge which is stored on a capacitor:  $Q = C_{el} U$ , where  $C_{el}$  is the electrical capacity of the piezo [cf. (221) on page 131]. The first term in (238) is the charge generated by the force acting on the piezo. The factor in front of the force  $F_3$  is the piezoelectric modulus again. It is the same number as in (231), although the unit obviously is [C/N] instead of [m/V] in (231). However, both units are the same and it will be shown below that both numbers need to be equal because the piezoelectric effect is reversible.

If a piezo is strained perpendicular to its polarization direction, there is also a charge generated on the electrodes. Therefore, another term needs to be added in (237):

$$D_3 = d_{31} \sigma_1 + d_{32} \sigma_2 + d_{33} \sigma_3 + \varepsilon_0 \varepsilon_r E_{e1}. \quad (239)$$

$$\Rightarrow Q_3 = d_{31} \frac{L_P}{h_P} F_1 + d_{32} \frac{b_P}{h_P} F_2 + d_{33} F_3 + \frac{\varepsilon_0 \varepsilon_r A_P}{h_P} U. \quad (240)$$

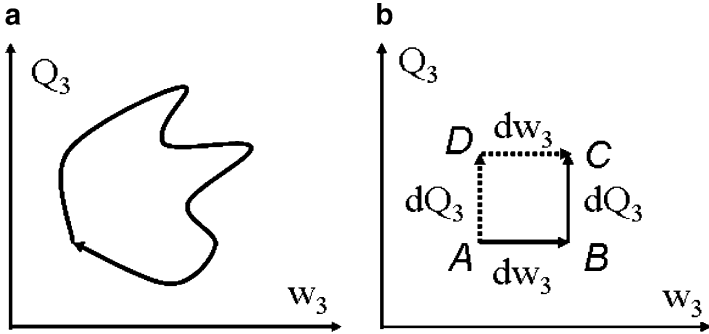
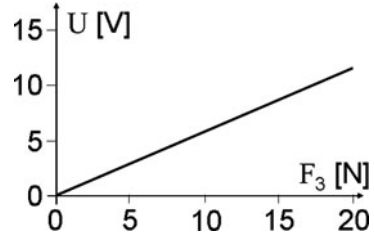
When a piezo is employed as a sensor, the above equation is used to calculate the charge generated by straining. Actually not the charge but the voltage is measured. The electrodes of the piezo are short-circuited, the contact between the electrodes is opened, and the force is applied. As the contact between the electrodes is open, no charge can flow onto or leave the electrodes; i.e.,  $Q_3$  is zero in (240) yielding the voltage  $U$  which can be measured:

$$\begin{aligned} U &= \frac{d_{31}}{\varepsilon_0 \varepsilon_r} \frac{L_P}{A_P} F_1 + \frac{d_{32}}{\varepsilon_0 \varepsilon_r} \frac{b_P}{A_P} F_2 + \frac{d_{33}}{\varepsilon_0 \varepsilon_r} \frac{h_P}{A_P} F_3 \\ &= g_{31} \frac{L_P}{A_P} F_1 + g_{32} \frac{b_P}{A_P} F_2 + g_{33} \frac{h_P}{A_P} F_3 \quad \text{with } g_{ij} = \frac{d_{ij}}{\varepsilon_0 \varepsilon_r}. \end{aligned} \quad (241)$$

The quantities  $g_{ij}$  are often also called piezoelectric moduli  $g_{31}$ ,  $g_{32}$ , and  $g_{33}$ . In this book, it is avoided to use these quantities to prevent confusion of the different types of piezoelectric moduli.

Equation (241) describes the characteristic curve of a piezo employed as a sensor. Figure 102 shows the characteristic curve of a piezo with thickness  $h_P$ , area  $A_P$ , Young's modulus  $E_P$ , piezoelectric modulus  $d_{33}$ , and relative permittivity  $\varepsilon_r$  of 0.5 mm, 10 mm<sup>2</sup>, 60 GPa,  $250 \times 10^{-12}$  m/V, and 2,400, respectively, calculated with (241). The force is acting in polarization direction and a few Newtons are enough

**Fig. 102** Characteristic curve of a piezo as a sensor



**Fig. 103** Paths in the charge-deflection plane

to generate an output of several volts. This is an extraordinary result, because other sensors, e.g., strain gauges, produce only some millivolts without amplification. Therefore, electromagnetic disturbances usually are no problem for piezoelectric sensors.

On the other hand, it needs to be noted that the piezoelectric modulus is a function of temperature and in almost every case temperature compensation is necessary. Besides this, every piezoelectric material has a small electrical conductivity resulting in a slow discharging of the piezo. Therefore, some electronics is necessary if long-time measurements are to be performed. Short-time measurements of changes of force or pressure are realized easier.

The piezoelectric modulus  $d_{33}$  in (231) and (238) is the same. This is due to the fact that the potential energy of a piezo is reversible [48]. That is, we start from a certain state where the piezo has a deflection  $w_i$  and charge  $Q_i$  resulting in a potential energy  $V_{p0}$ . Then, deflection and charge are changed such that after a while the piezo arrives at the same state (deflection and charge) again (cf. Fig. 103a). The energy of a piezo is the same as before the changes, if the potential energy is reversible. In other words, the potential energy is a function of deflection and charge only, and not a function of the way how this state was approached. It is obvious that the potential energy of a piezo is reversible, because it is a function of deflection and charge only. It is the sum of the potential energy of a capacitor and a spring:

$$V_p = \frac{1}{2} \frac{Q_3^2}{C_{el}} + \frac{k_3}{2} w_3^2. \tag{242}$$

Starting from the above equation, the piezo is deflected by  $\Delta w_3$  (path from *A* to *B* in Fig. 103b) resulting in the following energy:

$$V_p = V_{p0} + \frac{\partial V_p}{\partial w_3} \Delta w_3. \quad (243)$$

Then the charge of the piezo is changed by  $\Delta Q_3$  (path from *B* to *C* in Fig. 103b):

$$V_p = V_{p0} + \frac{\partial V_p}{\partial w_3} \Delta w_3 + \frac{\partial V_p}{\partial Q_3} \Delta Q_3 + \frac{\partial^2 V_p}{\partial w_3 \partial Q_3} \Delta w_3 \Delta Q_3. \quad (244)$$

If the charge is changed first and then the deflection (path from *A* over *D* to *C* in Fig. 103b), the energy is calculated as:

$$V_p = V_{p0} + \frac{\partial V_p}{\partial Q_3} \Delta Q_3 + \frac{\partial V_p}{\partial w_3} \Delta w_3 + \frac{\partial^2 V_p}{\partial Q_3 \partial w_3} \Delta Q_3 \Delta w_3. \quad (245)$$

If the potential energy of the piezo is reversible, (244) and (245) need to be equal, because the same state is approached via different ways (first changing the deflection and then the charge or vice versa):

$$\frac{\partial^2 V_p}{\partial Q_3 \partial w_3} = \frac{\partial^2 V_p}{\partial w_3 \partial Q_3}. \quad (246)$$

Differentiation of (242) yields:

$$\frac{\partial V_p}{\partial Q_3} = \frac{Q_3}{C_{el}} = U \text{ and } \frac{\partial V_p}{\partial w_3} = k_3 w_3 = F_3. \quad (247)$$

Thus, (246) results in:

$$\frac{\partial U}{\partial w_3} = \frac{\partial F_3}{\partial Q_3}. \quad (248)$$

Solving (231) for *U* and calculating the differential with respect to  $w_3$  on the one hand, and solving (238) for  $F_3$  and calculating the differential with respect to  $Q_3$  yields:

$$U = \frac{w_3 - F_3/k_3}{d_{33}} \Rightarrow \frac{\partial U}{\partial w_3} = \frac{1}{d_{33}}, \quad (249)$$

$$F_3 = \frac{Q_3 - C_{el}U}{d_{33}^*} \Rightarrow \frac{\partial F_3}{\partial Q_3} = \frac{1}{d_{33}^*}. \quad (250)$$

The asterisk in (250) denotes the piezoelectric modulus appearing in (238). According to (248), (249) and (250) are equal and thus  $d_{33}$  and  $d_{33}^*$  are equal.



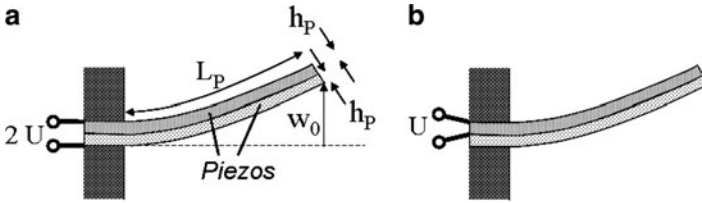


Fig. 104 Piezo bimorphs

The deflection of a piezo typically is on the order of several nanometers. This is not enough for most applications. On the other hand, the forces generated by piezos are comparatively large. Therefore, it is advantageous to partly sacrifice the force generated and to increase the deflection. This is typically achieved by an arrangement of two or more piezos bonded on each other. Figure 104a shows two piezos bonded onto each other. This is called a *piezo bimorph*. If a voltage is applied in polarization direction of a piezo, it shrinks laterally. Thus, the other piezo of the bimorph needs to be supplied with a voltage against its polarization direction resulting in a lateral extension. As a consequence, the bimorph is deflected transversally.

If voltage is applied to both piezos of a bimorph, the voltage is limited because always one of the piezos is charged against its polarization direction. Therefore, it is usual to apply power only to the piezo which shall shrink laterally (Fig. 104b). The deflection is reduced by a factor of 2, because only one piezo is active but the voltage may be raised more, e.g., a factor of 5. Thus, charging only one piezo yields larger deflections. Typically, the transversal deflection of the bimorph is a factor of ten larger than the lateral deflection of the two piezos.

The deflection  $w_{0,B}$  of a bimorph composed of two piezos with length  $L_P$ , width  $b_P$ , thickness  $h_P$ , Young's modulus  $E_P$ , and piezoelectric modulus  $d_{33}$  is calculated with the following equation [49] (Only one piezo is active.):

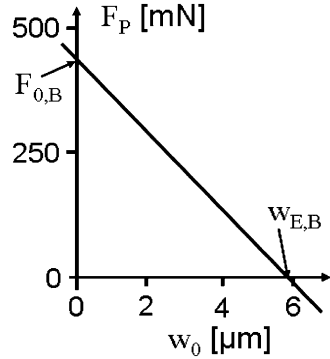
$$w_{0,B} = w_{E,B} + w_{F,B} = \frac{3}{8} d_{31} \frac{L_P^2}{h_P^2} U + \frac{L_P^3}{2 E_P b_P h_P^3} F. \quad (251)$$

The second term in (251) describes the deflection  $w_{F,B}$  of a beam with a rectangular cross-section and thickness  $2h_P$  which is clamped at one end and loaded transversally with a force  $F$  at the free end. This is already known from the theory of the deflection of beams [cf. Fig. 45 and (94) on page 65 and 66]. The first term in (251) describes the deflection  $w_{E,B}$  of the beam generated by the voltage  $U$ .

The force  $F_P$  generated by a piezo bimorph is calculated by solving (251) for  $F$  and changing the sign, because the force of the bimorph is calculated instead of the force acting on it:

$$F_P = \frac{3}{4} d_{31} \frac{E_P b_P h_P}{L_P} U - 2 \frac{E_P b_P h_P^3}{L_P^3} w_0. \quad (252)$$

**Fig. 105** Characteristic curve of a piezo bimorph as an actuator



Equation (252) describes the characteristic curve of a piezo bimorph. Figure 105 shows this curve for a bimorph with piezos with length, width, thickness, Young’s modulus, and piezoelectric modulus of 20 mm, 5 mm, 1 mm, 60 GPa, and  $190 \times 10^{-12}$  m/V, respectively, and driven by 200 V. The comparison with Fig. 101 on page 140 shows that a larger deflection and a smaller force are achieved. The characteristic curve is linear and subtends ordinate and abscissa at the maximum force  $F_{0,B}$  obtained when the deflection of the bimorph is prevented by outer forces and the maximum deflection  $w_{E,B}$  which is obtained when no outer forces are acting on the bimorph. The characteristic curve extends over the intersections, because a larger force than  $F_{0,B}$  acting against the bimorph bends it backward and an additional force in forward direction results in a larger deflection than  $w_{E,B}$  and a negative counter force generated by the elastic forces of the beam.

The maximum force  $F_{0,B}$  is the first term in (252), which calculates the force of the bimorph when it is not deflected. The maximum deflection is found as the first term of (251), which calculates the deflection when no force is acting on the bimorph:

$$F_{0,B} = \frac{3}{4} d_{31} \frac{E_P b_P h_P}{L_P} U, \tag{253}$$

$$w_{E,B} = \frac{3}{8} d_{31} \frac{L_P^2}{h_P^2} U. \tag{254}$$

The two above equations show that the characteristic curve of a piezo can be adapted in certain limits to the needs of a given application. A larger ratio of the thickness to the length increases the maximum force and reduces the maximum deflection. With a larger width, the maximum force can be enhanced without changing the maximum deflection.

If the bimorph is deflected in only one direction, the inactive piezo can be exchanged by a non-piezoelectric material. This opens up the door for optimizing deflection and force by varying Young’s modulus and thickness of the inactive layer.

The deflection  $w_0$  of a *piezo bonded onto an inactive carrier* consists of two terms describing the deflection due to the electric field  $w_E$  and due to external forces  $w_F$  similar as in the case of a bimorph (251) [49]:

$$w_0 = f_{w_E} w_{E,B} + f_{w_F} w_{F,B} =: w_E + w_F. \quad (255)$$

Here,  $w_{E,B}$  and  $w_{F,B}$  are the deflections of a bimorph due to the piezoelectric effect and outer forces as in (251), and  $f_{w_E}$  and  $f_{w_F}$  are correction factors which describe the influence of the carrier layer:

$$f_{w_E} = \frac{2}{(1 + h_V) + \left( (1 - E_V h_V)^2 / (4 E_V (1 + h_V)) \right)}, \quad (256)$$

$$f_{w_F} = \frac{8(1 + E_V)}{4 E_V (1 + h_V)^2 + (1 - E_V h_V)^2}. \quad (257)$$

In the two equations above, the quantities  $h_V$  and  $E_V$  appear which are defined as ratios of the thickness of carrier  $h_T$  and piezo  $h_P$  and their Young's moduli  $E_T$  and  $E_P$ , respectively:

$$h_V = \frac{h_T}{h_P} \quad \text{and} \quad E_V = \frac{h_T E_T}{h_P E_P}. \quad (258)$$

The maximum force  $F_0$  generated by the piezo on an inactive carrier is obtained by solving (255) for the force (contained in  $w_{F,B}$ ) and assuming zero deflection  $w_0$ :

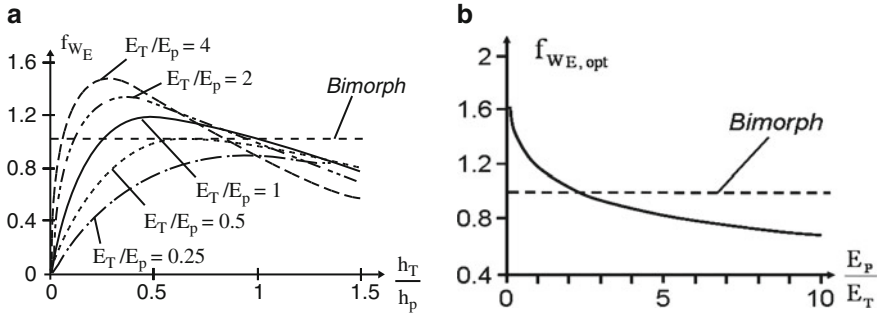
$$F_0 = \frac{E_V(1+h_V)}{(1+E_V)} \frac{3}{4} d_{31} \frac{E_P b_P h_P}{L_P} U = f_{F_0} F_{0,B} \quad \text{with} \quad f_{F_0} := \frac{E_V(1+h_V)}{(1+E_V)}. \quad (259)$$

The first fraction in the above equation is defined as the correction factor  $f_{F_0}$ , which describes the effect of bonding the piezo to a carrier beam, while the rest of the equation is the maximum force  $F_{0,B}$  of a bimorph as shown in (253).

The characteristic curve of a piezo on an inactive carrier is obtained by calculating the force  $F_P$  from (255) for the force and not assuming zero deflection:

$$F_P = f_{F_0} F_{0,B} - \frac{2 E_P b_P h_P^3}{f_{w_E} L_P^3} w_0. \quad (260)$$

Now the question rises what is the optimum thickness and Young's modulus to obtain maximum deflection  $w_E$  and maximum force  $F_0$ , respectively. According to (255), the maximum deflection  $w_E$  is calculated as the product of the maximum deflection  $w_{E,B}$  of the bimorph and the correction factor  $f_{w_E}$  which describes the effect of the carrier layer. Therefore, the maximum deflection is achieved with the parameters for which the correction factor is largest. Figure 106a shows the



**Fig. 106** Correction factor  $f_{w_E}$  of a piezo bonded to an inactive carrier calculated with (256) as a function of the ratio of the thicknesses of carrier and piezo (a), and value obtained with the thickness ratio optimum for large deflection (b)

correction factor  $f_{w_E}$  calculated with (256) as a function of the thickness ratio of carrier  $E_T$  and piezo  $E_P$  and the ratio of the Young's moduli. The dashed straight line corresponds to a bimorph. This shows that the bimorph is not optimum with respect to a large deflection. 20% more deflection can be achieved, if the thickness of the carrier is only half of the thickness of the piezo. The larger the ratio of the Young's moduli of carrier and piezo is, the larger the deflections can be obtained, if the thickness ratio of carrier and piezo is optimum.

The optimum thickness ratio is found by calculating the maximum of  $f_{w_E}$  as a function of this ratio. Thus, the derivative of  $f_{w_E}$  with respect to the thickness ratio is calculated, set to be zero, and solved for the thickness ratio. This calculation is hard to be done by hand but easily performed with the help of suitable computer codes:

$$\frac{\partial f_{w_E}}{\partial (h_T/h_P)} = 0 \Rightarrow \left(\frac{h_T}{h_P}\right)_{opt} = \frac{1}{2} \sqrt{\frac{E_P}{E_T}}. \quad (261)$$

If the above equation is inserted into (256), the maximum correction factor  $f_{w_E,opt}$  can be calculated which is obtained with the optimum thickness ratio:

$$f_{w_E,opt} = 32 \frac{2 + \sqrt{E_P/E_T}}{32 + 41\sqrt{E_P/E_T} + 8 E_P/E_T}. \quad (262)$$

The maximum correction factor  $f_{w_E,opt}$  calculated with the above equation is shown in Fig. 106b and clearly shows that larger deflections can be obtained when the carrier is made of a material with a larger Young's modulus. However, it needs to be considered that a different material for the carrier results in a thermal expansion different from the one of the piezo. Therefore, the deflection will change with temperature. Employing a carrier with a larger Young's modulus is only an advantage, if the cross sensitivity to temperature changes does not perturb the desired performance.

If not the deflection but the force generated by a piezo on a carrier shall be maximum, the correction factor  $f_{F_0}$  needs to be maximized. Figure 107 shows  $f_{F_0}$  as a function of the ratios of thicknesses and Young's moduli of carrier and piezo. The larger the thickness and the Young's modulus of the carrier are, the larger the forces are generated. There are no optimum values. However, the possible deflection will vary also as a function of these parameters, and a piezo on a carrier is normally not only chosen to obtain the maximum force, but the maximum deflection.

A piezo *bimorph* may be equipped with an inactive carrier between the piezos. This arrangement is shown in Fig. 108. This also corresponds to two piezos glued onto each other to form a bimorph. The glue behaves as an inactive layer, and the following equations show what is the effect of the glue on deflection and force of such a bimorph with an inactive carrier between the piezos.

In the following, the calculations assume that only one piezo is powered. In most applications, however, it will be advantageous to move the beam up by driving the upper piezo and down by the other one. That way, larger deflections are achieved.

The deflection of a bimorph with an inactive carrier between the piezos is calculated with (255) with modified correction factors [49]:

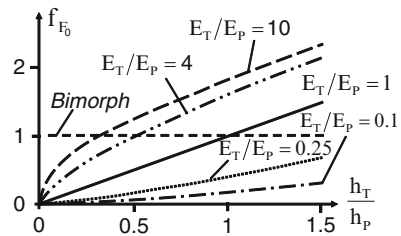
$$f_{w_E} = \frac{(1 + h_V)}{1 + (3/2)h_V + (3/4 + (1/8)E_V)h_V^2} \tag{263}$$

$$f_{w_F} = \frac{1}{1 + (3/2)h_V + (3/4 + (1/8)E_V)h_V^2}. \tag{264}$$

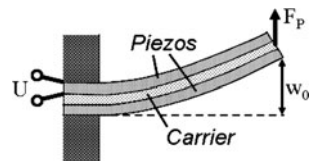
The maximum force and force as a function of the deflection (characteristic curve) are calculated with (259) and (260) but with a different correction factor:

$$f_{F_0} = (1 + h_V). \tag{265}$$

**Fig. 107** Correction factor  $f_{F_0}$  for a piezo bonded on an inactive carrier calculated with (259)



**Fig. 108** Bimorph with an inactive carrier between two piezos



If the thickness and Young’s modulus of the carrier shall be optimized to obtain either maximum deflection or maximum force, the correction factors according to (263) and (265) need to become maximum. These correction factors are drawn in Fig. 109 as a function of the ratios of the thicknesses and Young’s moduli of carrier and piezos, respectively.

The thicker the carrier and the larger its Young’s modulus are the smaller is the deflection. If the deflection shall be large, a carrier between the piezos is no advantage and the glue between the piezos of a bimorph should be as thin as possible and its Young’s modulus should be as small as possible.

On the other hand, the maximum force  $F_0$  generated by a bimorph with a carrier between the piezos is rising as a linear function of the thickness of the carrier but is not a function of its Young’s modulus.

Piezos and carriers do not need to be beams with a rectangular cross-section. Deflections and forces of *piezos and carriers as circular plates* simply supported at their rim (cf. Fig. 110) are calculated also with (255) and (259), if maximum deflection  $w_{E,B}$ , deflection due to outer forces  $w_{F,B}$ , and maximum force  $F_{0,B}$  of the bimorph are calculated with the following equations [49]:

$$w_{E,B} = \frac{3}{8} d_{31} \frac{R_p^2}{h_p^2} U, \tag{266}$$

$$w_{F,B} = 0.069 \frac{R_p^2}{E_p h_p^3} F, \tag{267}$$

$$F_{0,B} = 5.45 d_{31} E_p h_p U. \tag{268}$$

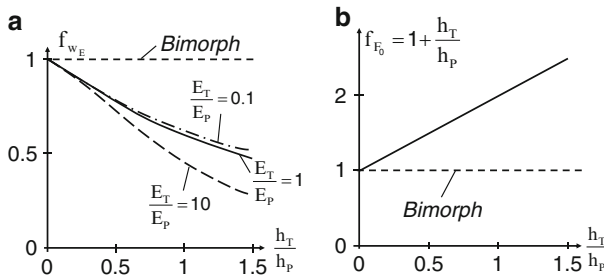
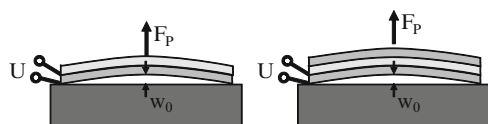


Fig. 109 Correction factors  $f_{w_E}$  and  $f_{F_0}$  for bimorphs with an inactive carrier in between as a function of the ratios of the thicknesses and Young’s moduli of carrier and piezo

Fig. 110 Circular piezo plates simply supported at their rim



In the above equations, it is assumed that Poisson's ratio both of piezo and carrier layer are 0.3. If arrangements of piezos and carriers are used, the same correction factors  $f_{w_E}$ ,  $f_{w_0}$ , and  $f_{F_0}$  are to be used as in (256), (257), and (259) for a piezo on a carrier or (263), (264), and (265) for two piezos with a carrier in between, respectively. This means also that the optimum ratios of thicknesses and Young's moduli of piezo and carrier are the same as described above for rectangular beams.

The characteristic curves of all piezos and arrangements of piezos with inactive carriers employed as actuators are straight lines defined by their maximum force  $F_0$  and maximum deflection  $w_E$  as shown in Fig. 111a. These parameters are listed in Table 13 and the required correction factors are found in Table 14. Thus, the characteristic curve of any arrangement of piezos discussed here is described by the following equation and can be calculated with the help of the two tables:

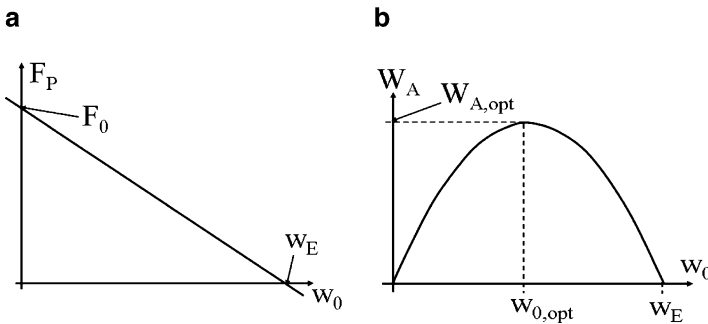
$$F_P = F_0 \left( 1 - \frac{w_0}{w_E} \right). \quad (269)$$

If the force which has to be overcome is not a function of the deflection of the piezo arrangement, e.g., a weight is to be lifted; the mechanical energy  $W_A$  generated by the piezo is the product of force and deflection:

$$W_A = F_P w_0 = F_0 w_0 \left( 1 - \frac{w_0}{w_E} \right). \quad (270)$$

This means that for the maximum force no work is done, because the deflection is zero, and for no force acting on the piezo also no work is done, although the deflection is maximum. Figure 111b shows the work  $W_A$  (= energy output) as a function of the deflection  $w_0$ . The maximum work  $W_{A,opt}$  can be calculated from the maximum of the parabola described by (270):

$$W_{A,opt} = \frac{1}{4} F_0 w_E. \quad (271)$$



**Fig. 111** (a) Characteristic curve of a piezo or an arrangement of piezos with a carrier and (b) work done by the piezo

The optimum deflection  $w_{0,opt}$ , where the work is maximum, is found at half of the maximum deflection  $w_E$ . Half of the maximum deflection corresponds also to half of the maximum force, because the characteristic curve in Fig. 111a is a strait line. The energy input is approximately the same independent of the deflection; it is the electrical energy  $W_C$  necessary to charge the capacity of the piezo [cf. (221) on page 133]. To obtain the exact energy input, the work done by the actuator would need to be added, but the energy required for charging is at least ten times larger.

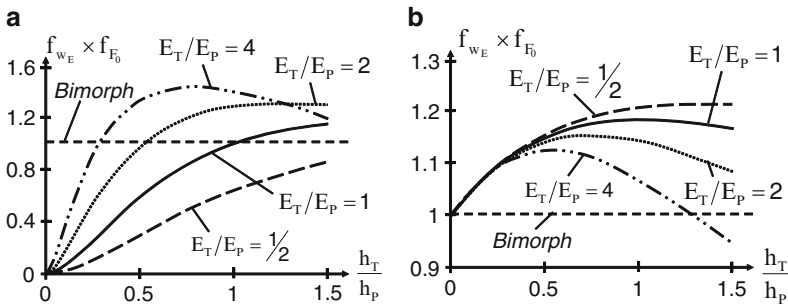
So, if a piezo is designed for a certain application and the energy consumption is an issue, e.g., in battery powered devices, the piezo should be dimensioned such that it is able to achieve the double of the desired deflection and force. Both a smaller and a larger piezo will result in a small efficiency.

The efficiency  $\eta_A$  is calculated from the ratio of the energy output  $W_{A,opt}$  and the energy input  $W_C$ . It is shown for every arrangement of piezos in Table 13. The maximum efficiency is on the order of 10%. Sometimes people say that the efficiency of piezos would be approximately 95%, but this is only true when the electrical energy stored during deflection of the piezo is recovered when the deflection is reduced again. Normally, this is not done and the 95% efficiency is only a theoretical value, while in practical applications less than 10% are achieved.

Thickness and Young’s modulus of carriers can also be optimized for maximum efficiency. As seen in Table 13, the product of the correction factors  $f_{w_E}$  and  $f_{F_0}$  needs to be optimized to achieve this. This product is shown in Fig. 112 both for a piezo on a carrier and a bimorph with a carrier between the piezos as a function of the ratios of the thicknesses and the Young’s moduli of piezo and carrier.

For a piezo on a carrier, the product is:

$$f_{w_E} f_{F_0} = \frac{2 E_V}{\left(1 + (1 - E_V h_V)^2 / (4 E_V (1 + h_V)^2)\right) (1 + E_V)} \tag{272}$$



**Fig. 112** Product of the correction factors  $f_{w_E}$  and  $f_{F_0}$  which need to be optimized to obtain maximum efficiency. (a) Piezo with a carrier, (b) bimorph with a carrier between the piezos



For a bimorph with a carrier between the piezos, the product is:

$$f_{wE} f_{F_0} = \frac{(1 + h_V)^2}{1 + (3/2)h_V + (3/4 + (1/8) E_V)h_V^2} \tag{273}$$

Figure 112a shows that for a piezo on a carrier, the energy output and efficiency can be largest when Young’s modulus of the carrier is maximum. A bimorph is not optimum with respect to efficiency; 20% more energy output could be achieved with the same input, if the inactive piezo would be two times thicker.

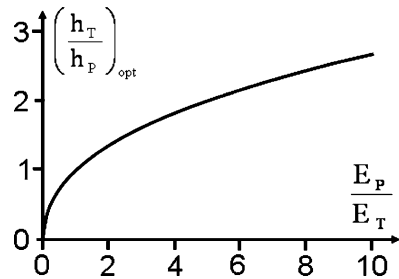
A bimorph with a carrier between the piezos achieves a higher efficiency when the carrier is made of a material with a small Young’s modulus and adapted thickness. The optimum thickness ratios is found by calculating the derivative of (273), setting it equal to zero, and solving for the thickness ratio. This calculation is very lengthy and yields a very complex result, but it can easily be obtained with suitable computer codes:

$$\left(\frac{h_T}{h_P}\right)_{opt} = \sqrt[3]{\frac{2 - E_T/E_P + 2\sqrt{1 - E_T/E_P}}{E_T/E_P}} + \sqrt[3]{\frac{E_T/E_P}{2 - E_T/E_P + 2\sqrt{1 - E_T/E_P}}} - 1. \tag{274}$$

The above equation is not easy to understand but when drawn in a graph shows its simple nature (cf. Fig. 113).

If the largest possible deflection is to be achieved, a piezo on a carrier is the best possible arrangement. The deflection can be enhanced by choosing a carrier with a large Young’s modulus. However, it needs to be considered that the deflection will be a function of temperature also, if the carrier is not a piezo (see “Thermal Actuators”). An alternative for obtaining a larger deflection with a suitable carrier layer is often to employ a larger bimorph. The thickness of the carrier should be close to the ratio given in (261) on page 147.

If the largest possible force is to be achieved, the best option is a piezo without any carrier pushing in polarization direction.



**Fig. 113** Optimum thickness ratio of carrier and piezo for maximum efficiency of a bimorph with a carrier between the piezos

Energy output and efficiency are largest for a piezo without any carrier layer. However, if at a considerably large deflection energy output and efficiency are to be maximized, a bimorph with an inactive carrier between the piezos often will be the best solution. Young’s modulus of the carrier should be as small as possible and its thickness should be designed according to (274). Besides the fact that the piezoelectric moduli are a function of temperature (see below), deflection and force of this arrangement will not be sensitive to temperature changes, because it is symmetrical with respect to the neutral fiber.

The maximum energy output density of a piezo actuator is calculated from (271) together with (236) (page 140) and (231) (page 139):

$$\frac{W_{A,opt}}{V} = \frac{1}{4} d_{33}^2 E_p \frac{U^2}{h_p^2} \tag{275}$$

If for maximum voltage, thickness, Young’s modulus, and piezoelectric modulus 400 V, 0.5 mm, 60 GPa, and  $250 \times 10^{-12}$  V/m are assumed, respectively,  $0.6 \mu\text{J}/\mu\text{L}$  is obtained for the maximum output energy density.

The piezoelectric moduli, e.g.,  $d_{33}$  and  $d_{31}$ , are a function of temperature (cf. Fig. 114). As a consequence, both sensors and actuators are cross sensitive to temperature changes, and this fact needs to be taken into account, when devices based of the piezoelectric effect are designed.

Besides the fact that the piezoelectric moduli are a function of temperature, there is the *pyroelectric effect*: When the temperature changes, electrical charges are generated on the electrodes of a piezo, and these charges cause deflection of the piezo. The charge change  $\Delta Q_3$  generated on the electrodes in polarization direction by a temperature change  $\Delta T$  is calculated with the following equation:

$$\Delta Q_3 = A_p p_3 \Delta T. \tag{276}$$

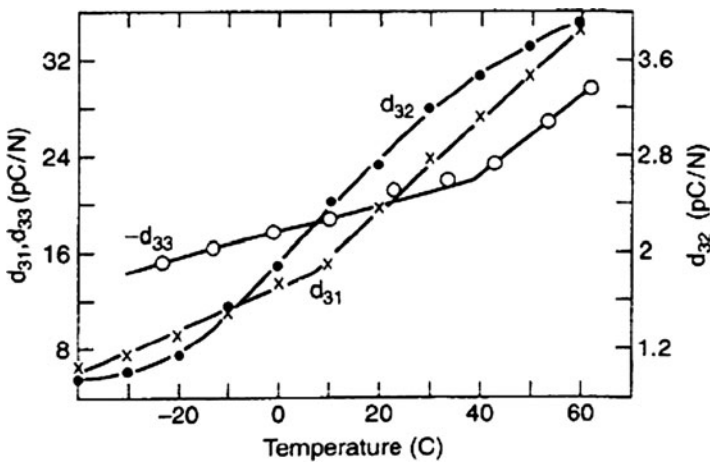


Fig. 114 Piezoelectric moduli of PVDF as a function of temperature [50]

In the above equation,  $p_3$  is the pyroelectric coefficient in the direction of polarization which is a material property of the piezo. The charge change  $\Delta Q_3$  corresponds to a voltage change  $\Delta U$  on the capacitor according to the definition of the electric capacity  $C_{el}$ :

$$\Delta Q_3 = C_{el} \Delta U = \varepsilon_0 \varepsilon_r \frac{A_P}{h_P} \Delta U = A_P p_3 \Delta T \Rightarrow \Delta U = \frac{h_P p_3}{\varepsilon_0 \varepsilon_r} \Delta T. \quad (277)$$

In principle, the pyroelectric effect could be employed to measure temperature changes or to actuate a piezo. However, the slow discharging of piezos is a disadvantage for sensor applications and there are a lot of alternatives for temperature measurement. The actuation of a piezo by a voltage is much quicker than by the pyroelectric effect. Therefore, the pyroelectric effect mostly causes cross-sensitivity and is not employed in any way.

For every case in Table 13, the deflection is calculated as  $w_0 = w_E + w_F$ . The correction factors describing the effect of the arrangement of the piezos are given in Table 14.

## Exercises

### *Problem 21*

Your chief tells you to design a microvalve which employs piezo actuators. It shall be a two-way valve with an orifice which is closed by a piezo beam from two combined layers from PZT (cf. Fig. E20). This way, temperature changes do not affect switching.

Only the upper piezo layer shall be driven, while the lower one is short-circuited. This way, it is possible to apply the whole available voltage of up to 200 V and no piezo gets depolarized. The diameter of the orifice is 200  $\mu\text{m}$ . The piezo material is commercially available with a width of 500  $\mu\text{m}$ . The thickness of each of the two layers is 500  $\mu\text{m}$ .

The distance between piezo beam and valve seat of an open valve needs to be a fourth of the diameter of the inlet and outlet to avoid that the flow resistance is a function of the gap between the beam and the valve seat.

- What length do you choose for the piezo beam to allow a sufficient opening of the valve?
- Up to what pressure difference can the valve be opened, if the piezo beam is designed with a length of 30 mm?
- What would be the maximum pressure difference against which the valve could work, if the active piezo would not be combined with an inactive one, but would press against the valve seat as shown in Fig. E21?
- Do you think that the design of Fig. E21 is feasible?

**Table 13** Equations for the calculation of deflections, forces, work outputs, and efficiencies of piezos

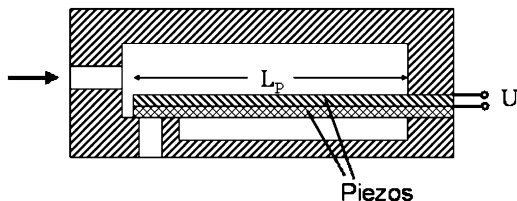
	Parallel to polarization	Perpendicular to polarization	Rectangular piezos with carrier	Circular plate with carrier
$w_E$	$d_{33} U$	$\frac{L_p U}{d_{31} h_p}$	$\frac{3}{8} d_{31} \frac{L_p^2 U}{h_p^2}$	$\frac{3}{8} R_p^2 d_{31} \frac{U}{h_p^2}$
$w_F$	$\frac{h_p F}{E_p A_p}$	$\frac{L_p F}{E_p b_p h_p}$	$\frac{L_p^3 F}{2 E_p b_p h_p^3}$	$0.069 \frac{R_p^2 F}{E_p h_p^3}$
$F_0$	$d_{33} \frac{E_p A_p}{h_p} U$	$d_{31} E_p b_p U$	$\frac{3}{4} \frac{E_p b_p h_p U}{L_p}$	$5.45 d_{31} E_p h_p U$
$W_{A,opt}$	$\frac{A_p E_p U^2 d_{33}^2}{4 h_p}$	$\frac{A_p E_p U^2 d_{31}^2}{4 h_p}$	$\frac{9 A_p E_p U^2 d_{31}^2}{32 h_p}$	$2.04 \frac{R_p^2 E_p d_{31}^2}{4 h_p}$
$\eta_A$	$\frac{E_p d_{33}^2}{2 \epsilon_0 \epsilon_r}$	$\frac{E_p d_{31}^2}{2 \epsilon_0 \epsilon_r}$	$\frac{9 E_p d_{31}^2}{64 \epsilon_0 \epsilon_r}$	$1.02 \frac{E_p d_{31}^2}{\epsilon_0 \epsilon_r}$

For the circular plates, a Poisson's ratio of 0.3 is assumed both for piezo and carrier

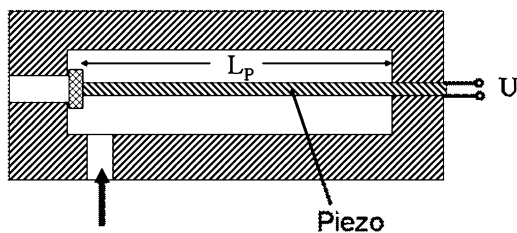
Here  $h_V = \frac{h_T}{h_P}$  and  $E_V = \frac{E_T}{E_P}$ .

**Table 14** Correction factors describing the effect of the arrangement of the piezos in Table 13

	Bimorph	Piezo with carrier	Two piezos on a carrier
$f_{wE}$	1	$\frac{2}{(1 + h_V) + (1 - E_V h_V)^2 / (4 E_V (1 + h_V))}$	$\frac{(1 + h_V)}{1 + (3/2)h_V + (3/4 + (1/8)E_V)h_V^2}$
$f_{wF}$	1	$\frac{8(1 + E_V)}{4 E_V (1 + h_V)^2 + (1 - E_V h_V)^2}$	$\frac{1}{1 + (3/2)h_V + (3/4 + (1/8)E_V)h_V^2}$
$f_{F0}$	1	$\frac{E_V (1 + h_V)}{(1 + E_V)}$	$(1 + h_V)$



**Fig. E20** Schematic view of the two-way valve



**Fig. E21** Cross-section of the alternative two-way valve

Piezoelectric modulus $d_{31}$ of PZT	$-170 \times 10^{-12} \text{ C/N}$
Piezoelectric modulus $d_{33}$ of PZT	$250 \times 10^{-12} \text{ C/N}$
Young's modulus of PZT	60 GPa
Relative permittivity of PZT	2,400

**Problem 22**

The 30-mm long beam of Fig. E21 shall now be used as a sensor.

- (a) What voltage can be measured when the beam is compressed with a force of 13 mN?
- (b) How much is the beam compressed by the load of 13 mN if the electrodes are short-circuited?
- (c) How much is the beam compressed by the load of 13 mN if the electrodes are not short-circuited but insulated to the environment? Take into account that a charge is generated by the force which diminishes the compression.

- (d) How much are the combined piezos of Fig. E20 deflected if the force of 13 mN is acting at their free end perpendicular to the direction of the beam and only one of the piezos is short-circuited and the other one has electrodes insulated to the environment?

### Problem 23

A piezo layer from PZT is bonded onto a circular silicon plate. The thickness of the piezo layer and the silicon plate are 0.5 and 1 mm, respectively, and the radius of both is 10 mm. The applied voltage is 200 V.

- (a) Calculate the maximum force and the maximum deflection of this arrangement. What is the maximum energy output?
- (b) Calculate the electrical energy which is needed to deflect this actuator. Calculate the ratio of energy output to energy input and, thus, calculate the efficiency. Hint: The piezo actuator forms a capacitor together with the electrodes. The energy stored in the capacitor is calculated with (221) on page 131.
- (c) Calculate the efficiency also for the case that the piezo is not combined with the silicon carrier and is working only by extension in the direction of polarization.
- (d) Calculate the efficiency for the case that the lower side of the silicon is bonded to another piezo layer of the same size as on the upper side. There is always voltage applied to only one piezo layer.
- (e) What is your conclusion, when you compare the results of (b)–(d)?

Young's modulus of PZT	60 GPa	Young's modulus of silicon	190 GPa
$d_{31}$ of the PZT layer	$-170 \times 10^{-12}$ C/N	Relative permittivity of PZT	2,400
$d_{33}$ of the PZT layer	$250 \times 10^{-12}$ C/N		

# Thermal Actuators

If the linear dimensions of a device are reduced, its mass decreases with the third power, while its surface decreases only proportional to the square (cf. page 3). As a consequence, the ratio of surface to volume or mass is very large for microdevices. This means, that a microdevice is heated up much more quickly and with less energy consumption, because of its small mass, and is cooling down more quickly due to its large surface to mass ratio. Therefore, heating up is a suitable actuation principle for a lot of microdevices, while it is not a good solution for macroscopic devices. The ink-jet printer is a good example (cf. page 251).

If an isotropic rigid body with a coefficient of thermal extension  $\alpha_{th}$  is heated up by the temperature change  $\Delta T$ , it extends to all directions by the strain  $\epsilon_{th}$  (cf. Fig. 115):

$$\epsilon_{th} = \alpha_{th} \Delta T. \tag{278}$$

The coefficient of thermal extension  $\alpha_{th}$  is a material constant which needs to be measured or found in a suitable book.

If a pressure load  $p$  is acting on the heated body against the direction of thermal extension, the strain generated by the pressure according to Hooke's law needs to be added:

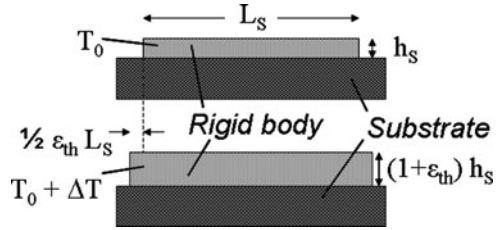
$$\epsilon_{th} = \alpha_{th} \Delta T - \frac{p}{E_S}. \tag{279}$$

In the above equation,  $E_S$  is Young's modulus of the heated body. The deflections  $w_x$  and  $w_z$  of the rigid body in direction of its length  $L_S$  and thickness  $h_S$ , respectively, under the action of compressive forces  $F_x$  and  $F_z$  are derived from the above equation ( $b_S$  is the width of the body):

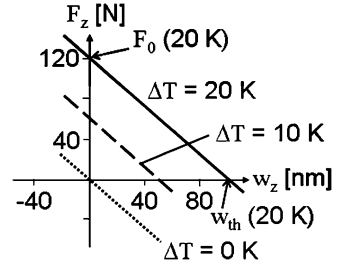
$$w_z = \alpha_{th} h_S \Delta T - \frac{h_S}{L_S b_S} \frac{F_z}{E_S} + \frac{\nu_S}{b_S} \frac{F_x}{E_S}, \tag{280}$$

$$w_x = \alpha_{th} L_S \Delta T - \frac{L_S}{h_S b_S} \frac{F_x}{E_S} + \frac{\nu_S}{b_S} \frac{F_z}{E_S}. \tag{281}$$

**Fig. 115** Extension of a heated rigid body



**Fig. 116** Characteristic curves of a thermal actuator as calculated with (282)



From the two above equations, the characteristic curves of this thermal actuator can be calculated by solving for the force. For example, if deflection and force are acting in z-direction only, it is obtained:

$$F_z = L_S b_S E_S \left( \alpha_{th} \Delta T - \frac{w_z}{h_S} \right). \tag{282}$$

Figure 116 shows the characteristic curve of a thermal actuator with length, width, and thickness of 5 mm, 2 mm, 0.5 mm, respectively. A Young’s modulus of 60 GPa is assumed, so that this actuator corresponds to the piezo described by the characteristic curve in Fig. 101 on page 140. The comparison of Figs. 116 and 101 shows that a thermal actuator is able to produce similar forces and deflections as a piezo when it is heated up by just 10 K, and it can be heated up much more. Therefore, it needs to be noted that thermomechanical actuators are capable to produce large forces and small deflections.

Similar as for piezos, the maximum force  $F_0$  and maximum deflection  $w_{th}$  without any force acting on the actuator are calculated from (282) and (280), respectively:

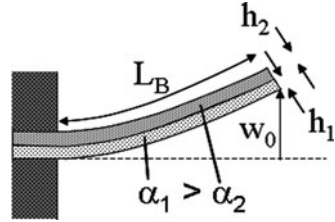
$$F_0 = L_S b_S E_S \alpha_{th} \Delta T. \tag{283}$$

$$w_{th} = h_S \alpha_{th} \Delta T. \tag{284}$$

As a consequence of small deflections and large forces generated by a thermo-mechanical actuator, as in the case of piezo actuators, it is advantageous to employ



Fig. 117 Bimaterial actuator



beams with different deflection which are bonded to each other. Such beams are called *bimaterial actuators*. They consist of two stripes of materials with different thermal extension. Figure 117 shows such a bimaterial actuator. The thermal extension of beam 1 is larger than the one of beam 2. That is why the beams are bending upward. Obviously, the coefficient of thermal expansion of beam 1 should be as large as possible and the one of beam 2 as small as possible or even negative. There are some ceramic materials which show a negative coefficient of thermal extension and, therefore, shrink when heated.

When heated, both beams of a bimaterial actuator undergo a change in their length according to (281) (with no forces acting). The difference in their straining takes the role which was played by the straining of a piezo when a voltage is applied. Thus, the same equations can be employed as for bimorphs when the piezoelectric strain is replaced by the difference in the thermal strain of the two beams:

$$\varepsilon_1 = d_{31} \frac{U}{h_p} \rightarrow \varepsilon_{th} = \Delta\alpha_{th} \Delta T. \tag{285}$$

This way, the deflections and characteristic curves of a bimaterial actuator are obtained from (255) (on page 146) and (259)/(260):

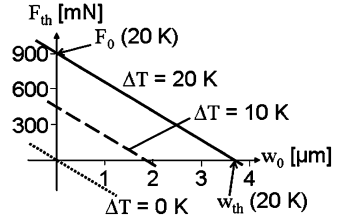
$$w_0 = w_{th} + w_F = f_{wE} \frac{3 L_B^2}{4 h_2} \Delta\alpha \Delta T + f_{wF} \frac{1}{2} \frac{L_B^3}{E_2 b_B h_2^3} F, \tag{286}$$

$$F_{th} = f_{F0} \frac{3 E_2 b_B h_2^2}{2 L_B} \Delta\alpha \Delta T - \frac{2 E_2 b_B h_2^3}{f_{wF} L_B^3} w_0. \tag{287}$$

In the above two equations, the correction factors  $f_{wE}$ ,  $f_{wF}$ , and  $f_{F0}$  are the same as in (256), (257) and (259) (see page 146), respectively,  $h_1$  and  $E_1$  are thickness and Young’s modulus of the beam with the larger thermal expansion, and  $b_B$  and  $L_B$  are width and length of the beams, respectively.

The characteristic curve of a bimaterial actuator consisting of two beams with the same dimensions and Young’s modulus as in Fig. 116 and with a difference in the coefficients of thermal expansion  $\Delta\alpha = 5 \times 10^{-6}$  is shown in Fig. 118 as calculated with (287). The characteristic curve is a strait line as in Figs. 116 and 105 (page 145). As a consequence, the optimum deflection for maximum energy

**Fig. 118** Characteristic curves of a bimaterial actuator as calculated with (287)



output and efficiency is half of the maximum deflection  $w_{th}$  and the maximum energy is a fourth of the product of maximum deflection  $w_{th}$  and maximum force  $F_0$  [cf. (271) on page 150].

If the dimensions of a bimaterial actuator shall be optimized for maximum deflection, force, or efficiency, the correction factors  $f_{w_E}$ ,  $f_{F_0}$ , and their product need to be maximized as in the case of the piezo actuators. Therefore, the same ratios of thickness and Young’s modulus are optimum as described on pages 146f and 152f, respectively.

The energy input  $W_E$  necessary to deflect a thermomechanical actuator approximately is the heat necessary to enhance its temperature, because this is much larger than the energy output which can be generated. An actuator with mass  $m_K$  and heat capacity  $C_{th}$  which is to be heated by the temperature difference  $\Delta T$  consumes the following power:

$$W_E = m_K C_{th} \Delta T. \tag{288}$$

The efficiency of a bimaterial actuator is calculated as the ratio of energy output to input. The maximum efficiency typically is on the order of 0.1%. In Table 15, the efficiency is shown together with deflection, force, and maximum energy output of thermomechanical actuators. The correction factors are the same as for piezos with a carrier which are listed in Table 14 on page 156. There is no thermomechanical equivalence to two piezos with a carrier in between, because it is virtually impossible to heat only one side of a stack of beams bonded to each other.

Maximum force is achieved with a plate in normal direction. Maximum deflection is obtained with a bimaterial beam.

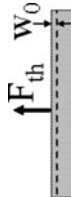
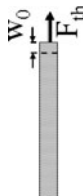
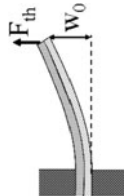
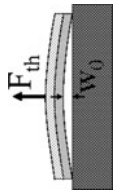
Table 15 shows also that a certain deflection or force can be achieved (in certain limits) either by enlarging the actuator or by raising the temperature at which it is driven. When the efficiency shall not become too small, it is better to drive a small actuator at large temperatures. A larger temperature will also help to extenuate two other principles of thermal actuators: cross sensitivity to changes of ambient temperature and long cooling time.

The maximum energy output density of the actuator is calculated from the optimum output energy in Table 15 divided by the volume  $h_S A_S$ . If for Young’s modulus, thermal expansion, and temperature difference  $120 \times 10^9$ ,  $15 \times 10^{-6}$ , and  $200^\circ\text{C}$  are assumed, respectively, a maximum energy output density of  $270 \mu\text{J}/\mu\text{L}$  is obtained.

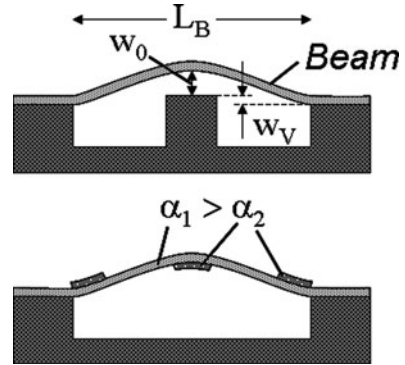
**Table 15** Equations for the calculation of deflections, forces, work outputs, and efficiencies of thermomechanical actuators

	In normal direction, one plate	In lateral direction, one plate	Two rectangular beams	Circular bimaterial plates
$w_{th}$	$h_S \alpha_{th} \Delta T$	$L_B \alpha_{th} \Delta T$	$\frac{3 L_B^2}{8 h_2} \Delta \alpha_{th} \Delta T$	$\frac{3 R_p^2}{8 h_2} \Delta \alpha_{th} \Delta T$
$w_F$	$\frac{h_p}{E_S A_S} F$	$\frac{L_S}{E_S b_S h_S} F$	$\frac{L_B^3}{2 E_2 b_B h_2^3} F$	$\frac{R_p^2}{E_2 h_2^3} F$
$F_0$	$A_S E_S \alpha_{th} \Delta T$	$h_S b_S E_S \alpha_{th} \Delta T$	$\frac{3 E_2 b_B h_2^2}{4 L_B} \alpha_{th} \Delta T$	$f_{F_0} 5.45 E_2 h_2^2 \Delta \alpha_{th} \Delta T$
$W_{A,opt}$	$\frac{h_S A_S}{4} E_S \alpha_{th}^2 \Delta T^2$	$\frac{h_S L_S b_S}{4} E_S \alpha_{th}^2 \Delta T^2$	$f_{w_E} f_{F_0} \frac{9}{64} h_2 L_B b_B E_2 \Delta \alpha^2 \Delta T^2$	$f_{w_E} f_{F_0} 0.51 h_2 R_p^2 E_2 \Delta \alpha^2 \Delta T^2$
$\eta_A$	$\frac{h_S A_S}{4 m_K C_{th}} \alpha_{th}^2 \Delta T$	$\frac{h_S L_S b_S}{4 m_K C_{th}} \alpha_{th}^2 \Delta T$	$f_{w_E} f_{F_0} \frac{9 h_2 L_B b_B E_2 \Delta \alpha^2 \Delta T}{64 m_1 C_{th,1} + m_2 C_{th,2}}$	$f_{w_E} f_{F_0} 0.51 \frac{h_2 R_p^2 E_2 \Delta \alpha^2 \Delta T}{m_1 C_{th,1} + m_2 C_{th,2}}$

For the circular plates, a Poisson's ratio of 0.3 is assumed for both layers of the bimaterial actuator



**Fig. 119** Beam clamped at both ends deflected by heating only the beam



Another kind of thermomechanical actuator may be built of a *heated beam clamped at both ends*. If only the beam and not its bearing is heated, the beam generates compressive stress and buckles when its critical stress is exceeded. Typically, the beam is heated by an electrical current through a part of the beam or a conductor path on it. Some predeflection  $w_V$  of the beam is necessary to ensure buckling to the desired side (cf. Fig. 119 top). Another possibility is a material with a different heat expansion at least on a part of the beam forming a bimaterial. Figure 119 bottom shows some possibilities. Even a layer of the same material as the membrane on one of its sides will result in a buckling towards this side, because the neutral fiber is shifted towards that side. It needs also to be taken into account that a thin film may deflect a beam due to initial stress as shown in Figs. 7 or 9 on page 12 and 13, respectively.

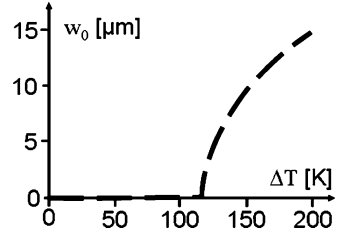
When heated, a beam not clamped and without any outer forces acting on it, expands according to (278) (page 159). When the beam is clamped at both ends its stress becomes more compressive by the thermal stress  $\sigma_{th}$  which is the strain without clamping ( $\alpha_{th} \Delta T$ ) times Young's modulus  $E_B$  of the beam. When the buckling stress is overcome, the beam buckles. The deflection as a function of the stress of the beam is calculated with (116) (page 75). Inserting  $\sigma_{th}$  into (116) yields:

$$w_0(F = 0) = 3\sqrt{2}\sqrt{\frac{I}{A_B}}\sqrt{\frac{\sigma_0 - E_B \alpha_{th} \Delta T}{\sigma_k} - 1}. \quad (289)$$

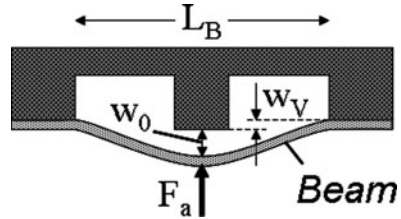
Figure 120 shows the deflection  $w_0$  of a rectangular beam with Young's modulus  $E_B = 100$  GPa, thermal expansion  $\alpha_{th} = 5 \times 10^{-6}$ , thickness  $20 \mu\text{m}$ , width  $60 \mu\text{m}$ , length  $1$  mm calculated with the above equation as a function of temperature. For temperatures where the stress of the beam is less than the critical stress  $\sigma_k$ , there is no deflection. When the stress exceeds the critical stress, the deflection quickly rises as calculated with (289).

The interrelationship between the force  $F$  acting perpendicular to the center of the beam and its deflection  $w_0$  is calculated by inserting the thermal stress  $\sigma_{th}$  into

**Fig. 120** Deflection of a heated beam clamped at both ends as calculated with (289)



**Fig. 121** Heated beam working against an external force  $F_a$



(114) (on page 74). The force  $F_B$  generated by the beam is obtained from this by changing the sign:

$$F_B = -4 \frac{A_B}{L_B} w_0 \left( 48 \frac{E_B I}{A_B L_B^2} + \sigma_0 - E_B \alpha_{th} \Delta T + \frac{8}{3} E_B \frac{w_0^2}{L_B^2} \right). \quad (290)$$

The force  $F_B$  generated by the heated beam is shown in Fig. 122 as a function of the deflection  $w_0$  for several temperature changes  $\Delta T$ . The curves have been calculated with (290) and the parameters used for Fig. 120. The temperature changes correspond to different stresses in the beam, and, therefore, Fig. 122 shows a similar graph as Fig. 54 on page 75.

If a beam as shown in Fig. 121 is pressed by an external force  $F_a$  onto the structure ensuring the predeflection  $w_V$ , heating of the beam will cause no deflects until the curve shown in Fig. 122a reaches  $F_a$ . In Fig. 122, this is achieved at 100 K. The black filled circles denote the force and deflection of the beam at certain temperature changes  $\Delta T$ . When the temperature exceeds 100 K, the beam deflects as indicated in the figure.

If the external force  $F_a$  is larger, the beam achieves this force at a larger temperature (150 K in Fig. 122b). At this temperature, the beam suddenly snaps up from 5  $\mu\text{m}$  to approximately 11  $\mu\text{m}$  as indicated by the arrow. This is an effect which might be employed in a microsystem to achieve a certain effect. For example, a valve or an electrical switch could be opened quickly by a certain amount. If this is not desired, the predeflection  $w_V$  needs to be designed larger than the snapping over deflection  $w_U$  which can be calculated with (117) (on page 75).

It is also seen in Fig. 122b that a predeflection  $w_{V,2}$  which is designed too small or no predeflection results in no deflection of the beam even at very high temperatures.

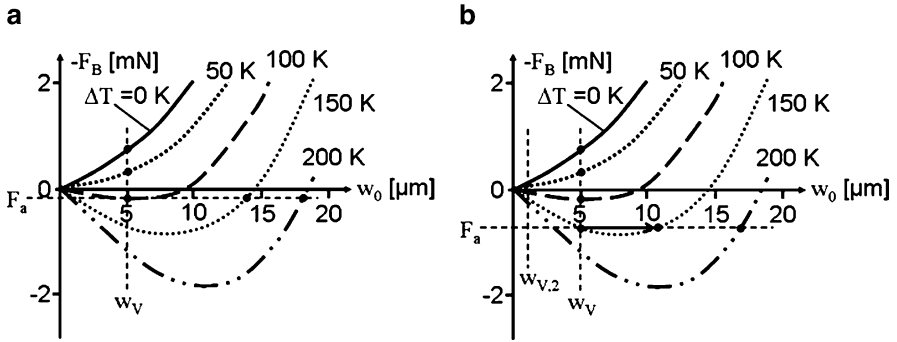


Fig. 122 Heated beam working against an external force  $F_a$

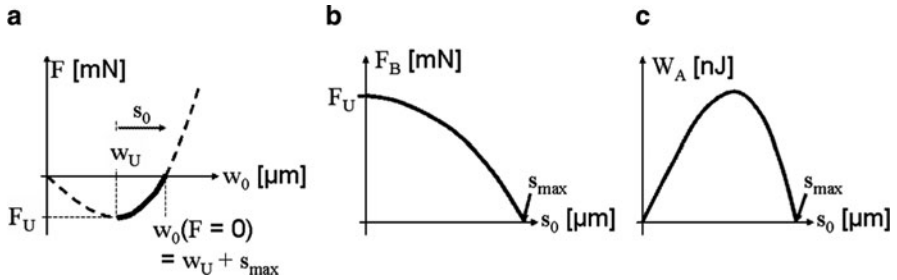


Fig. 123 Redefinition of the deflection  $s$  of the heated beam (a) and characteristic curve (b). (c) The energy output

The reason why the beam does not deflect without any predeflection is, that it is at its unstable equilibrium where only small or theoretically no force is necessary to hold it in position.

The deflection of the heated beam starts at its predeflection  $w_V$ . Therefore, it appears to be appropriate to define a deflection  $s_0$  starting at  $w_V$  and to redraw the characteristic curve of the heated beam clamped at both ends as an actuator (cf. Fig. 123a, b). Maximum force can be achieved, if for a given temperature the predeflection is designed equal to the deflection  $w_U$  of the snapping over of the beam resulting in the characteristic curve shown in Fig. 123b.

The characteristic curve is calculated by substituting  $w_0$  in (290) by  $(w_U + s_0)$ :

$$F_B = 4 \frac{A_B}{L_B} [w_U + s_0] \left( 48 \frac{E_B I}{A_B L_B^2} + \sigma_0 - E_B \alpha_{th} \Delta T + \frac{8}{3} E_B \frac{[w_U + s_0]^2}{L_B^2} \right). \quad (291)$$

The maximum force  $F_U$  of the actuator is either found from (118) (on page 76) when the thermal stress is added to the residual stress or from (291) by inserting (117) (page 75) for  $w_U$  and setting  $s_0$  to zero:

$$F_U = 8\sqrt{\frac{2}{3}} \frac{\sqrt{A_B I}}{L_B} \sigma_k \left( \frac{\sigma_0 - \alpha_{th} E_B \Delta T}{\sigma_k} - 1 \right)^{\frac{3}{2}} \tag{292}$$

The maximum deflection  $s_{max}$  is the difference between the predeflection  $w_U$  and the deflection without any external force acting (cf. Fig. 123a). It is found by subtracting (117) from (116) and adding the thermal stress to the residual one:

$$s_{max} = 3\sqrt{2} \left( 1 - \frac{1}{\sqrt{3}} \right) \sqrt{\frac{I}{A_B}} \sqrt{\frac{\sigma_0 - \alpha_{th} E_B \Delta T}{\sigma_k} - 1}. \tag{293}$$

The energy output  $W_A$  of the actuator when loaded with a force which is independent of the deflection is calculated as the product of the force  $F_B$  (291) and the deflection  $s_0$ :

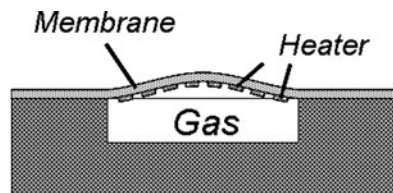
$$W_A = F_B s_0. \tag{294}$$

The energy output is drawn in Fig. 123c as a function of the deflection. The comparison of the characteristic curve and the energy output of a heated beam clamped at both ends (Fig. 123), and a bimaterial actuator (Fig. 118 on page 162) shows only small differences. However, the bimaterial actuator is not sensitive to changes of the stress due to outer forces. The tolerance required to achieve the suitable predeflection of a heated beam clamped at both ends may also be very small.

Another kind of thermal actuator is the *thermo-pneumatic actuator*. A thermo-pneumatic actuator consists of a chamber closed by a membrane and some means to heat the gas (mostly air) inside of the chamber. An example is shown in Fig. 124 where a heater coil on the membrane is heated up. As a consequence, the heated air is expanding and generates a pressure drop  $\Delta p$  over the membrane. The pressure difference bulges up the membrane to a deflection  $w_0$ . The behavior of a closed gas volume is described quite well by the ideal gas law:

$$p V = n_{mol} R_G T = \frac{V_0}{V_{mol}} R_G T. \tag{295}$$

In the above equation,  $p$ ,  $V$ , and  $T$  represent pressure, volume, and temperature of the gas, respectively, while  $n_{mol}$  and  $R_G = 8.31 \text{ J / (mol K)}$  denote the number of



**Fig. 124** Thermo-pneumatic actuator

moles of the gas contained in the volume and the gas constant which is a natural constant.  $V_0$  is the volume of the gas when the membrane is not deflected and  $V_{\text{mol}}$  is the volume of a mole of the gas at normal temperature and pressure (22.4 L at 101.3 kPa and 295 K). As an approximation, it is assumed here that the entire gas in the actuator chamber is heated to the same temperature. In general, this is not true and the temperature of the heater needs to be much larger than the mean temperature of the gas. Therefore, the temperature discussed here is a kind of effective temperature which generates the calculated effects.

When the gas in the actuator is heated and an outer force  $F_0$  prevents the deflection of the membrane, the volume inside the actuator chamber equals  $V_0$ . The absolute values of the outer force and the force generated by the actuator are equal to the ratio of the pressure drop  $\Delta p$  over the membrane and its area  $A_M$ . From (295) it is obtained:

$$\Delta p = \frac{F_0}{A_M} = \frac{R_G}{V_{\text{mol}}} \Delta T \Rightarrow F_0 = \frac{A_M R_G}{V_{\text{mol}}} \Delta T. \quad (296)$$

The above equation shows that the force generated at no deflection of the membrane is not a function of the volume  $V_0$  of the actuator.

The deflection of the membrane itself requires a certain pressure difference (cf. page 33ff). If an actuator is to be designed, the radius of the membrane will be chosen so large and its thickness so small that the pressure drop required to deflect the membrane itself is negligible compared with the effect of outer forces. Therefore, it appears to be reasonable to neglect the effect of the membrane on force and deflection in the following calculations. However, it needs to be checked in every case that this assumption is really true, and, if necessary the following equations need to be adapted.

If the deflection (i.e., volume change) and pressure difference of a thermo-pneumatic actuator is to be calculated under the action of outer forces in general, it helps to assume that the deflection was achieved in two steps. In the first step, the volume is extended by  $\Delta V$  by enhancing the temperature by  $\Delta T_1$ , while the pressure is kept constant at the value  $p_0$  when the membrane is not deflected. From (295):

$$\Delta V = \frac{V_0 R_G}{V_{\text{mol}} p_0} \Delta T_1 \Rightarrow \Delta T_1 = \frac{V_{\text{mol}} p_0}{V_0 R_G} \Delta V. \quad (297)$$

In the second step, the pressure is enhanced by a temperature change ( $\Delta T - \Delta T_1$ ) at a constant volume. Equation (295) now results in:

$$\begin{aligned} \Delta p &= \frac{V_0 R_G}{V_{\text{mol}}(V_0 + \Delta V)} (\Delta T - \Delta T_1) = \frac{V_0 R_G \Delta T - V_{\text{mol}} p_0 \Delta V}{V_{\text{mol}}(V_0 + \Delta V)} \\ &= \frac{R_G \Delta T - V_{\text{mol}} p_0 (\Delta V/V_0)}{V_{\text{mol}}(1 + (\Delta V/V_0))}. \end{aligned} \quad (298)$$

Equation (297) was inserted in (298) for  $\Delta T_1$ . Equation (298) is the characteristic curve of the thermo-pneumatic actuator. Figure 125a shows this characteristic



curve for an actuator with a volume  $V_0$  of 50 and 100 nL, respectively. The pressure  $p_0$  at no deflection and temperature enhancement  $\Delta T$  are 101 kPa and 100 K, respectively. It is seen that the maximum pressure  $\Delta p_{\max}$  at no deflection of the membrane is not a function of the volume  $V_0$  of the actuator. The characteristic curves extend over the coordinate axes, i.e., the maximum pressure  $\Delta p_{\max}$  and maximum volume change  $\Delta V_{\max}$  can be exceeded when an outer force or pressure pushes the membrane more down to negative values or pulls it more up. Maximum pressure and maximum volume change can easily be calculated by inserting  $\Delta V = 0$  and  $\Delta p = 0$  into (298), respectively:

$$\Delta p_{\max} = \frac{R_G \Delta T}{V_{\text{mol}}} \text{ and } \Delta V_{\max} = \frac{V_0 R_G \Delta T}{V_{\text{mol}} p_0}. \quad (299)$$

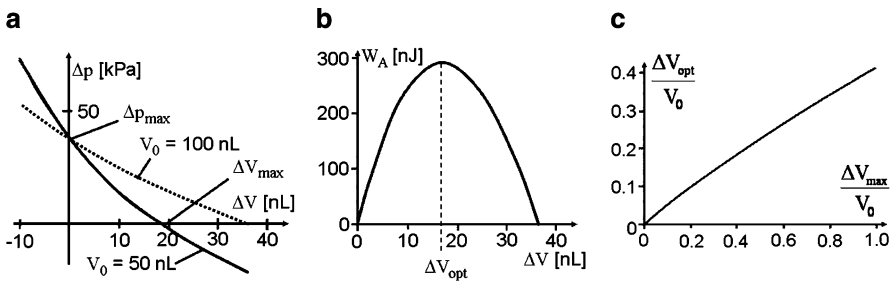
The energy output  $W_A$  of a thermo-pneumatic actuator is the product of the generated pressure difference  $\Delta p$  and volume change  $\Delta V$ . Thus from (298) it is obtained:

$$W_A = \frac{R_G \Delta T - V_{\text{mol}} p_0 (\Delta V / V_0)}{V_{\text{mol}} (1 + (\Delta V / V_0))} \Delta V. \quad (300)$$

The energy output of the actuator shown in Fig.125a, when working against a constant load is displayed in Fig.125b for an actuator volume  $V_0 = 100$  nL. As for a piezo, a thermal bimorph or a heated beam clamped at both ends, there is an optimum volume change  $\Delta V_{\text{opt}}$  (optimum deflection for the other actuators) at which the energy output is maximum. The efficiency at the maximum energy output is only on the order of 1% instead of approximately 10% for a piezo. This optimum volume change  $\Delta V_{\text{opt}}$  is found by calculating the zero of the derivative of (300):

$$\frac{\partial W_A}{\partial \Delta V} = 0 \Rightarrow \Delta V_{\text{opt}} = V_0 \left( \sqrt{1 + \frac{\Delta V_{\max}}{V_0}} - 1 \right) = V_0 \left( \sqrt{1 + \frac{R_G \Delta T}{V_{\text{mol}} p_0}} - 1 \right). \quad (301)$$

The volume change  $\Delta V_{\text{opt}}$  as calculated with the above equation is approximately half of the maximum volume change  $\Delta V_{\max}$ , if the  $\Delta V_{\max}$  is small compared



**Fig. 125** Characteristic curves of thermo-pneumatic actuators with a volume of 50 and 100 nL, respectively (a), energy output  $W_A$  (b), and optimum volume change for maximum energy output (c)

with the actuator volume without membrane deflection. Figure 125c shows the optimum volume change as a function of the maximum one calculated with (301).

The maximum energy output which can be generated is obtained now by inserting (301) for the volume change into (300):

$$W_{A,opt} = \frac{R_G \Delta T - V_{mol} p_0 (\sqrt{1 + (R_G \Delta T / (V_{mol} p_0))} - 1)}{V_{mol} \sqrt{1 + (R_G \Delta T / (V_{mol} p_0))}} V_0 \left( \sqrt{1 + \frac{R_G \Delta T}{V_{mol} p_0}} - 1 \right). \quad (302)$$

The maximum energy output density is obtained from the above equation by dividing by  $V_0$ . If 200°C and 100 kPa are assumed for the maximum possible temperature change and the environmental pressure, respectively, the maximum energy output density is 10  $\mu\text{J}/\mu\text{L}$ .

For most applications not only the volume change but also the deflection of the thermo-pneumatic actuator is important. The deflection of an actuator with a circular membrane can be calculated with sufficient accuracy, if a spherical cap is assumed for the form of the deflected membrane. The volume  $V_W$  of a spherical cap with height  $w_0$  (corresponding to the membrane deflection) and cap diameter  $2R_M$  (corresponding to the membrane diameter) is given by:

$$V_W = \frac{1}{6} \pi w_0 (3 R_M^2 + w_0^2) \approx \frac{1}{2} \pi w_0 R_M^2, \quad (303)$$

$V_W$  is the volume change generated by a temperature change. Therefore, from (297) the maximum membrane deflection  $w_{0,max}$  is obtained now:

$$V_W = \frac{V_0 R_G}{V_{mol} p_0} \Delta T \approx \frac{1}{2} \pi w_{0,max} R_M^2 \Rightarrow w_{0,max} = \frac{2}{\pi R_M^2} \frac{V_0 R_G}{V_{mol} p_0} \Delta T. \quad (304)$$

If a cylindrical actuator chamber with height  $h_A$  and radius  $R_M$  is assumed under the membrane, its volume is  $V_0 = h_A \pi R_M^2$ , and it is obvious that the maximum deflection of the membrane is not a function of the lateral dimensions but only of the height  $h_A$  of the actuator chamber:

$$w_{0,max} = \frac{2 h_A R_G}{V_{mol} p_0} \Delta T. \quad (305)$$

Similarly, the volume change  $V_W$  is calculated for square membranes. The volume change is found by integrating the deflection  $w_0$  as described by (65) on page 45 over the entire membrane:

$$V_W = \int_{-a_M/2}^{a_M/2} \int_{-a_M/2}^{a_M/2} w_0 \left( 1 - 4 \frac{x^2}{a_M^2} \right)^2 \left( 1 - 4 \frac{y^2}{a_M^2} \right)^2 dx dy = \left( \frac{8}{15} \right)^2 a_M^2 w_0. \quad (306)$$

In the same way as above, it is found now:

$$V_W = \frac{V_0 R_G}{V_{mol} p_0} \Delta T \approx \left(\frac{8}{15}\right)^2 w_{0,max} a_M^2 \Rightarrow w_{0,max} = \left(\frac{15}{8}\right)^2 \frac{V_0 R_G}{a_M^2 V_{mol} p_0} \Delta T. \quad (307)$$

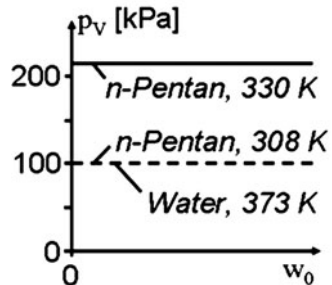
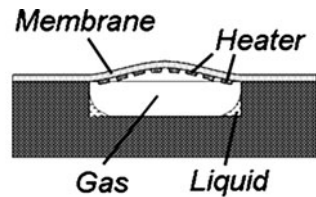
If an actuator chamber in the form of a hollow cuboid with height  $h_A$  is assumed, the maximum deflection is:

$$w_{0,max} = \left(\frac{15}{8}\right)^2 \frac{h_A R_G}{V_{mol} p_0} \Delta T. \quad (308)$$

Very large pressure changes can be generated by the *phase transition* of some material inside of an actuator chamber (see Fig.126 top). A simple example is a micro-steam engine: A small quantity of water is enclosed in an actuator chamber and the water is heated by some means until it is partly evaporated and the vapor pressure deflects the membrane. The vapor pressure  $p_V$  is only a function of temperature  $T$  and vaporization heat  $\Lambda$  of 1 mol:

$$p_V = e^{-\frac{\Lambda}{R_G T}}, \quad (309)$$

That is, the pressure generated is not a function of the deflection of the membrane as long as there are both liquid and gaseous water in the actuator. If the deflection of the membrane is changed by an outer force acting on the membrane, vapor is turned into liquid or vice versa and the pressure remains to be the same (cf. Fig.126 bottom).



**Fig. 126** Thermo-pneumatic actuator with phase transition (top) and characteristic curve (bottom)

A thermal actuator with phase transition can achieve comparatively large deflections and large forces even if the volume of the actuator chamber is small. If the actuator chamber is shallow and the diameter of the membrane is comparatively large, the deflection of the membrane is not a limiting factor for the volume expansion. Thus, a volume expansion of a factor of 10 appears to be possible. If from Fig. 126 a pressure of only 100 kPa is taken and multiplied by a volume change of 10, an energy output density of 1 MPa = 1,000  $\mu\text{J}/\mu\text{L}$  is obtained.

However, it is difficult to fill a small amount of liquid into the actuator chamber and to make sure that the liquid, and especially the vapor, does not leave it again by diffusion through the membrane.

The thermal energy  $W_{\text{th}}$  needed to generate deflection of the membrane consists of two terms. The first one is the energy necessary to heat up the liquid and the second one is the energy required for evaporation:

$$W_{\text{th}} = c_p \rho_F V_{\text{fl}} \Delta T + \lambda_v \rho_F \Delta V_{\text{fl}}. \quad (310)$$

Here  $c_p$ ,  $\rho_F$ ,  $\lambda_v$ ,  $V_{\text{fl}}$ , and  $\Delta T$  are the heat capacity, density, steam heat, volume, and temperature change of the liquid, respectively. The first term in the above equation is needed to generate the actuator pressure, while the second one generates the deflection of the membrane. In general, the first term is much larger than the second one. Therefore, in the light of minimizing the energy consumption the volume of the liquid inside the actuator needs to be chosen not larger than enough to allow for the desired maximum deflection.

The mass  $\Delta m$  of the liquid which needs to be evaporated to obtain a volume change  $\Delta V$  can be calculated from the mass  $m_{\text{mol}}$  and the volume  $V_{\text{mol}}$  of a mole of a gas:

$$\Delta m = \frac{\Delta V}{V_{\text{mol}}} m_{\text{mol}}. \quad (311)$$

The volume of a mole as a function of temperature and pressure is calculated from (295) (page 167) and the volume  $V_{\text{fl}}$  of a liquid is the ratio of its mass  $\Delta m$  and density  $\rho_F$ . Taking all this together with the above equation yields the minimum liquid volume  $V_{\text{fl,min}}$  required to generate a certain volume change  $\Delta V$ :

$$V_{\text{fl,min}} = \frac{\Delta V}{\rho_F} \frac{p}{R_G T} m_{\text{mol}}. \quad (312)$$

In general, the efficiency of all thermal actuators cannot be larger than the efficiency  $\eta_C$  of the Carnot process:

$$\eta_C = \frac{\Delta T}{T_0}. \quad (313)$$

In the above equation,  $T_0$  and  $\Delta T$  are the ambient temperature and the temperature rise generated for the process. As a consequence, it is advantageous to drive a thermal actuator high above the ambient temperature. Besides a higher efficiency, a higher driving temperature also reduces the effect of changes of the ambient

temperature on the performance of the actuator. In addition, this is necessary to cope with another typical disadvantage of thermal actuators: In general, it is easy to heat them up very quickly by an electrical current, but it is difficult to cool them down in a short time again. Cooling is achieved by heat diffusion in most cases and that is a comparatively slow process. Therefore, the heat capacity of thermal actuators needs to be designed as small as possible allowing quick heating and especially cooling.

In Table 16, the order of magnitude of typical properties of microactuators are compared with each other. The values are only an orientation for choosing the right actuator for a certain application and more exact results need to be calculated with the equations in the corresponding chapters.

## Exercises

### Problem 24

Figure E22 shows a microvalve which was introduced at the conference Actuator 96 by the company Bosch [51]. On a nearly circular silicon membrane, there are arranged two annular aluminum stripes. The valve is opened against pressure acting at the inlet by heating the inner aluminum ring. The geometry of the valve is designed such that it is closed when the inner ring is not heated.

- (a) What is the purpose of the outer aluminum ring? It is not heated to switch the valve.

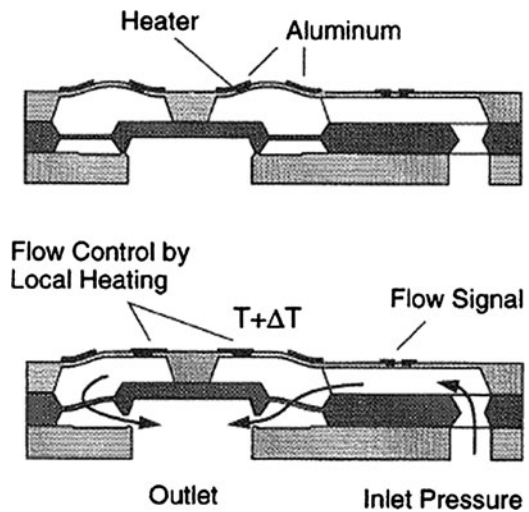
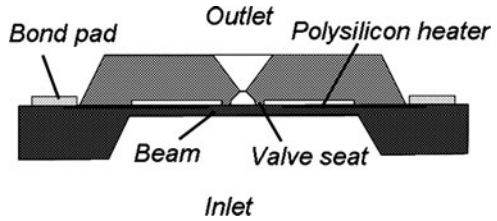


Fig. E22 Cross-section of the microvalve [51]



**Fig. E23** Cross-section of the microvalve [52]



- (b) It is difficult to calculate the behavior of an annular bimaterial. Therefore, rectangular bimaterial stripes shall be calculated here which are as wide as the circumference of the real aluminum and as long as their width. Calculate maximum force and maximum deflection of the bimaterial stripe. What is the maximum energy output?
- (c) Calculate the thermal energy needed to deflect the bimaterial stripe. Calculate the ratio to the energy output, and, thus find out the maximum efficiency. Hint: The energy required to deflect the stripe is approximately the same as the energy needed to heat it up.

Young's modulus of Al	70 GPa	Young's modulus of Si	190 GPa
Circumference of inner aluminum stripe	1 mm	Width of inner aluminum stripe	3.5 mm
Thickness of aluminum	6.7 $\mu\text{m}$	Thickness of membrane	12 $\mu\text{m}$
Thermal expansion of Al	$23.8 \times 10^{-6}/\text{K}$	Thermal expansion of Si	$2.7 \times 10^{-6}/\text{K}$
Heat capacity of Al	896 J/(kg K)	Heat capacity of Si	703 J/(kg K)
Density of aluminum	2,730 kg/m <sup>3</sup>	Density of silicon	2,330 kg/m <sup>3</sup>
Temperature increase of inner aluminum ring	200 K		

### Problem 25

On the conference Actuator 94, a microvalve with a silicon beam as a switching element was presented [52]. On the beam, there was a conductor path employed as a heater. The beam was deflected downward when heated and opened the valve (cf. Fig. E23).

When the beam is not heated, it lies horizontally on the valve seat and closes the valve (residual stress  $\sigma_0 = 0$ ). In the drawing, it is not shown that the beam is deflected a little bit downward also when the valve is closed, because the valve seat is a bit protruding.

- (a) The beam needs to move down for at least a quarter of the diameter of the outlet to avoid a limitation of the flow through the open valve. How much needs the beam to be heated to achieve the necessary deflection? (Assume that there is no

pressure drop over the beam when the valve is open and that, therefore, no force is acting on the beam).

- (b) What temperature change is necessary to open the valve against a pressure difference of 100 kPa?
- (c) What predeflection is necessary to allow opening of the valve with the temperature difference calculated at (b).
- (d) What is the deflection of a beam over the valve seat with the predeflection from (c) and the temperature difference from (b)?

Length of the beam	2.6 mm	Young's modulus of Si	190 GPa
Width of the beam	600 $\mu\text{m}$	Thermal expansion of Si	$2.3 \times 10^{-6}/\text{K}$
Thickness of the beam	20 $\mu\text{m}$	Residual stress of beam	0 MPa
Diameter of valve outlet	360 $\mu\text{m}$		

### ***Problem 26***

A thermo-pneumatic actuator with a cylindrical chamber which is built up as shown in Fig. 124 (page 167) shall close the inlet of a microvalve against an outer pressure. With (301) (page 169), the optimum stroke of the actuator can be calculated with which the energy is used most efficiently.

- (a) Please derive from the characteristic curve [(298) on page 168] an equation with which the counter pressure can be calculated at which the energy of the actuator is used optimally.
- (b) Which counter pressure is optimal with respect to energy consumption, if the actuator is heated by 100 and 200 K, respectively?
- (c) What is the necessary height of the actuator chamber if the stroke is 20  $\mu\text{m}$  at a temperature enhancement of 200 K and if the energy of the actuator is to be used optimally?
- (d) How large should the radius of the actuator be chosen to achieve the largest possible efficiency for the actuator?
- (e) How many energy is necessary at least to heat up the actuator by 200 K if the chamber is 30  $\mu\text{m}$  high and has a radius of 500  $\mu\text{m}$ ? Assume (unrealistically) that no heat is lost and only the air in the actuator is to be heated.
- (f) How large is the efficiency in this idealized case?

Heat capacity of air at constant pressure	1.005 kJ/(kg K)	Density of air	1.29 kg/m <sup>3</sup>
Pressure in the actuator when deflected without counter pressure			101.3 kPa



# Microoptics

Miniaturized optical components are very popular nowadays. For example, lenses in cellular phones need to be small and light weight and shall be equipped with an optical zoom. Micro-optical components are employed also in sensors, for data transmission, and chemical analysis.

Miniaturization is a disadvantage for optical imaging because every optical component is an aperture in the optical path and every aperture results in diffraction. As shown on the left of Fig. 127, light due to its wave nature does not propagate as straight lines but shows diffraction. The light intensity far behind a narrow aperture illuminated by parallel light does not show a rectangular profile. For the same reason, the focus of a lens with diameter  $D_L$  and focal length  $f_B$  cannot be arbitrarily small but the intensity distribution  $I_L$  is given by the following equation:

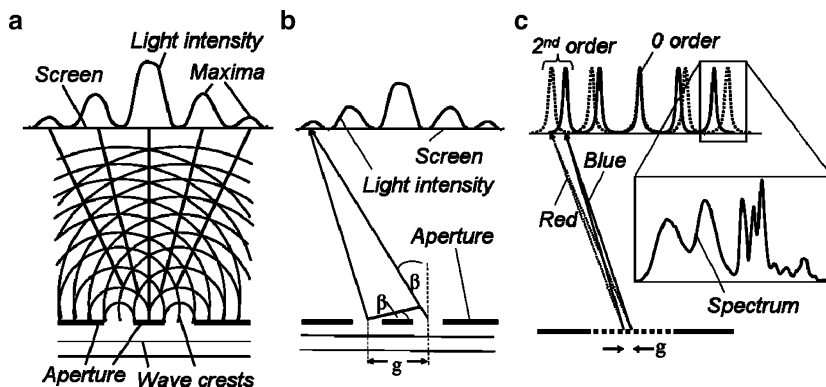
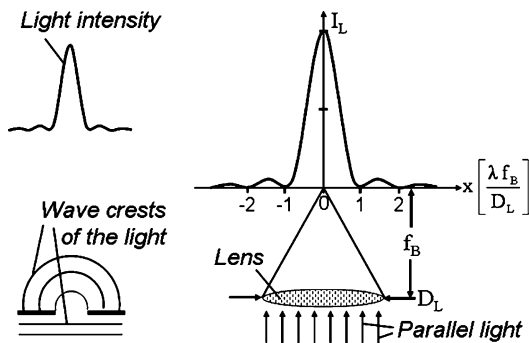
$$I_L = I_0 \left[ \frac{\sin(\pi x(D_L/\lambda f_B))}{(\pi x(D_L/\lambda f_B))} \right]^2. \quad (314)$$

In the above equation,  $\lambda$ ,  $x$ , and  $I_0$  are the wavelength of the light used, distance from the optical axis, and intensity of the light illuminating the aperture, respectively. The intensity distribution calculated with (314) is shown on the right of Fig. 127. Obviously, a smaller diameter  $D_L$  of the lens broadens the intensity distribution. As a consequence, an image projected with a lens becomes sharper when the lens diameter is enlarged and miniaturization of the lens results in a diffuse projection. Therefore, there are always other reasons for miniaturization of optical components which are more important for a certain application than the loss in sharpness, e.g., small mass, small size, and low cost.

If an optical spectrum is to be analyzed, usually a *diffraction grating* is employed. A diffraction grating is a large number of narrow optical apertures arranged next to each other. Typically, transparent or reflecting parallel slits are made on a surface which is opaque besides the slits.

To describe how an optical grating works, first only two transparent slits are considered – a so-called double slit. A plane wave arriving at the two slits can be considered behind the double slit as two point sources of light which are emitting

**Fig. 127** Optical diffraction of light waves at an aperture (left) and a lens (right)



**Fig. 128** Light interference behind a double slit (a, b) and a grating (c)

light in phase as spherical waves (cf. Fig. 128a). At certain angles, the wave crests and wave troughs of these waves coincide. On a screen far behind the double slit, the light intensity shows maxima at these angles. The formation of maxima is called constructive interference, the formation of minima where the crest of one wave compensates the trough of another wave is called destructive interference.

The angles  $\beta$  of maximum intensity are found where the optical path difference between the two waves is an integer multiple of the wavelength  $\lambda$ . Thus, as seen in Fig. 128b, intensity maxima are found where the following equation is fulfilled:

$$m \lambda = g \sin(\beta) \Rightarrow \sin(\beta) \approx \beta = \frac{m \lambda}{g}. \tag{315}$$

In the above equation,  $g$  is the distance of the centers of the two slits. It is called the grid constant.  $m$  is an integer which is called the diffraction order.

If not only two but also several or many slits with a constant distance  $g$  are employed, the maxima of all the spherical waves propagating from these slits interfere with each other and generate a grid of narrow maxima of equal intensity

on the screen (cf. Fig. 128c). Except the zero order maximum ( $m = 0$ ), the angle of constructive interference  $\beta$  is a function of the wavelength. The higher the order of a maximum is, the more is the spectrum spread and the larger is the resolution for analyzing it. As a consequence, a grid can be used to separate the light into its colors and to display and analyze its spectrum.

If the light is not illuminating the grid perpendicularly, the angle of incidence needs also to be taken into account when the position of the maxima are calculated. Figure 129 shows that the following equation needs to be satisfied for constructive interference:

$$m \lambda = g(\sin(\alpha) + \sin(\beta)) \Rightarrow \sin(\alpha) + \sin(\beta) = \frac{m \lambda}{g} \approx \alpha + \beta. \quad (316)$$

Note that  $\alpha$  and  $\beta$  as shown in Fig. 129 are defined to be positive. One of them can be negative when it is on the opposite side of the dashed line normal to the grid. By choosing  $\alpha$  the designer can control the position of, e.g., the first-order maximum of a certain wavelength.

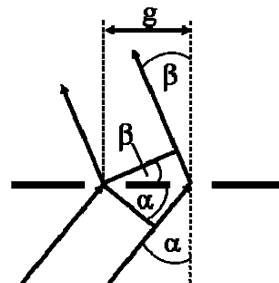
Besides transmission grids, there are also reflection grids. As shown in Fig. 130a, at a reflection grid, the reflected light is subject to interference. The angle  $\beta$  in (316) is negative in most cases of a reflection grid.

If a grid is to be used with low-intensity light, it is a disadvantage that roughly half of the light to be analyzed is absorbed. This problem is avoided when a *step grid* is employed as shown in Fig. 130b. The light reflected at a step is interfering with those reflected at the neighboring steps. As a result of the step, there is a phase difference between light reflected at different steps.

According to the reflection law, the angle  $\alpha_r$  of the reflected light is the same as the angle of incidence  $\alpha$  (cf. Fig. 131):

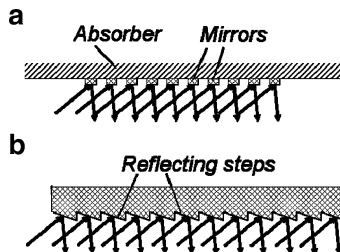
$$\alpha = \alpha_r. \quad (317)$$

Due to the slope angle  $\Phi_{St}$  of the steps, there is a wavelength for which the angles of reflection and constructive interference coincide. For this wavelength, the so-called *blaze wavelength*  $\lambda_B$ , and, to some smaller extent, the neighboring wavelengths

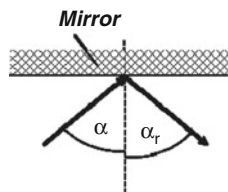


**Fig. 129** Interference at a not perpendicularly illuminated grid

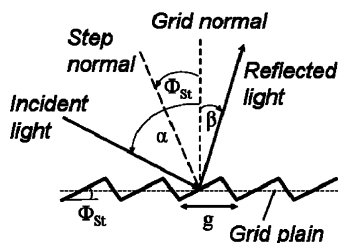
**Fig. 130** Reflection grid (a) and step grid (b)



**Fig. 131** Reflection at a mirror



**Fig. 132** Blaze angle at a step grid



the intensity of the interference maximum is enhanced. This effect can be employed for the design of spectrometers. If the detector is less sensitive in a certain range of wavelengths, e.g., extreme blue light, this can be compensated partly with a blaze in this range. Another possibility is that a certain range of the spectrum is of high importance for analytical purposes.

In Fig. 132, it is seen that the reflection law results in:

$$\alpha - \Phi_{St} = \Phi_{St} - \beta \Rightarrow \Phi_{St} = \frac{\alpha + \beta}{2}. \tag{318}$$

Note that  $\beta$  in Fig. 132 is negative according to the definition of (316). If  $\beta$  is the blaze angle for the blaze wavelength  $\lambda_B$ , both (318) and (316) need to be fulfilled, and the step angle  $\Phi_{St}$  is derived as:

$$\Phi_{St} = \frac{m \lambda_B}{2g}. \tag{319}$$

The blaze angle  $\beta$  at which constructive interference and reflection coincide is a function of the angle of incidence. It is found from the two above equations:

$$\beta = \frac{m \lambda_B}{g} - \alpha. \tag{320}$$

If the screen shown in Fig. 128b is far behind the double slit, the angles  $\beta$  of light emitted from the neighboring slits and interfering at a certain position on the screen are nearly the same. Therefore, (315) and (316) have been derived correctly. In microtechnique, however, the distance of the screen is never far from the grid. Therefore, the plain of the grid needs to be bent to achieve that light emitted at the same difference between incident angle  $\alpha$  and diffraction angle  $\beta$ , i.e., the same phase difference, arrives at the same point. This problem is solved when the light source (or entrance slit of a spectrometer), the grid, and the screen (or detector of a spectrometer) are arranged on a circle (cf. Fig. 133a). Such a circle is called a *Rowland circle*. It is a geometrical fact that every triangle inside of a given circumscribed circle and with one side identical to a certain secant shows the same angle opposite to the secant (cf. Fig. 133b).

The surfaces of optical components may show only small roughness, because even a small roughness will result in interference effects. Typically, the components need to be even within a margin which is a 20th of the wavelength.

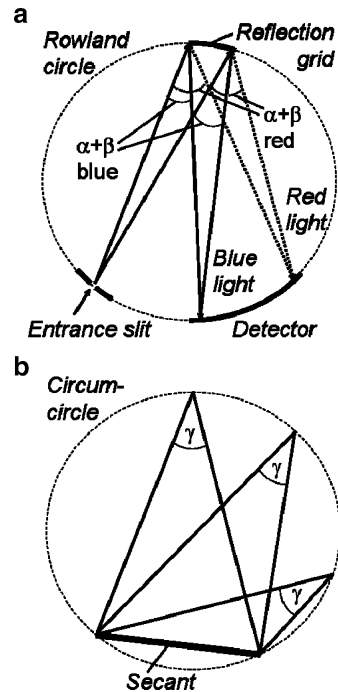


Fig. 133 Rowland circle

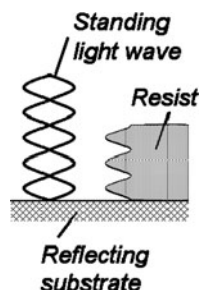
If light is arriving at a mirror at an angle of incidence of  $0^\circ$ , the incident wave is interfering with the reflected one. As a consequence, the so-called *standing wave* is generated and the intensity of the light in front of the mirror shows maxima and minima (cf. Fig. 134). At the surface of the mirror, there is a minimum and the distance from one minimum to the next one is half of the wavelength of the light. This phenomenon is observed when a partly transparent screen is placed very near to the surface of the mirror and moved back and forth to the mirror. Standing light waves are also the reason for interference patterns in photo resists on a reflecting substrate which cause uneven side walls of resist patterns after development.

An optical component suitable for miniaturization which is often used for analytical purposes is the *Fabry-Perot interferometer*. It consists of two mirrors which are arranged in parallel to each other (cf. Fig. 135). At least one of these mirrors is partly transparent. Therefore, light can enter into the space between the mirrors and a small part of this light is escaping out of the interferometer again. The integer multiples of the wavelengths  $\lambda$  of the light leaving the interferometer equal the double distance  $d$  between the mirrors:

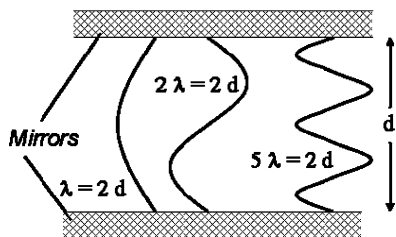
$$m \lambda = 2 d. \quad (321)$$

The reason for this effect is that the standing waves developing on the surfaces of the two mirrors can only interfere constructively if their minima and maxima coincide. That is, the standing wave between the mirrors needs to have a minimum on the surfaces of both mirrors.

**Fig. 134** Standing light wave on a reflecting substrate (*left*) and resist edge after developing (*right*)



**Fig. 135** Fabry-Perot interferometer and standing light waves developing inside



The path of light may also change its direction by *refraction*. Refraction occurs when the medium in which the light is propagating shows a change of its refractive index. The refractive index is the ratio of the velocity of light in vacuum and in the medium. So, the refractive index of the vacuum equals 1. In a medium, the velocity of light is a little bit smaller than in vacuum, and, therefore, the refractive index of media is a little bit larger than 1.

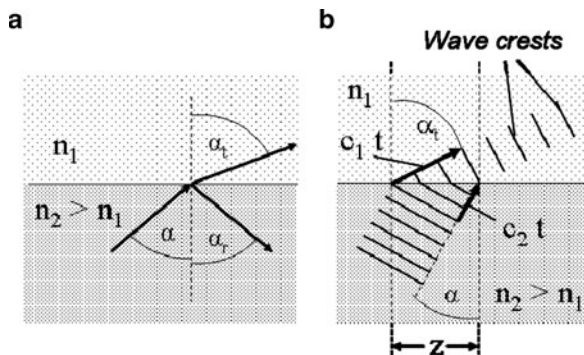
When light is propagating towards a down step in refraction index at an angle of incidence  $\alpha$  larger than  $0^\circ$ , it is partly reflected at an angle  $\alpha_r$  according to (317), and partly transmitted at an angle  $\alpha_t$  into the medium with the smaller refraction index (see Fig. 136a). One side of each wave crest will arrive earlier at the interface than the other (cf. left side of wave crest in Fig. 136b). The transmitted part of the side arriving earlier will continue propagating with a larger velocity  $c_1$ , while the other side for a small time  $t$  retains its former velocity  $c_2$ . The distances  $c_1 t$  and  $c_2 t$  the light is traveling within the time  $t$  are not equal. Therefore, the direction of propagation will change to a larger angle  $\alpha_t$ .

From the geometry in Fig. 136b, it is seen that:

$$c_1 t = z \cos(90^\circ - \alpha_t) = z \sin(\alpha_t) \text{ and } c_2 t = z \sin(\alpha)$$

$$\Rightarrow \frac{c_1 t}{\sin(\alpha_t)} = \frac{c_2 t}{\sin(\alpha)} \Rightarrow \frac{c_1}{c_2} = \frac{n_2}{n_1} = \frac{\sin(\alpha_t)}{\sin(\alpha)}. \tag{322}$$

Equation (322) is *Snell's law* with which the refraction angle can be calculated. Since  $n_2$  is larger than  $n_1$ , at a certain angle of incidence  $\alpha_{Tot}$ ,  $\alpha_t$  needs to become larger than  $90^\circ$ . This would mean that the light goes back into the medium where it comes from. As a consequence, for angles of incidence larger than  $\alpha_{Tot}$ , there is no refraction but only reflection according to the reflection law (317).  $\alpha_{Tot}$  is called the *angle of total reflection*. It can be calculated from (322) by choosing  $90^\circ$  for the angle  $\alpha_t$  of the transmitted light, i.e., the sine equals 1:



**Fig. 136** Reflection and refraction of light

$$\sin(\alpha_{\text{Tot}}) = \frac{n_1}{n_2} \Rightarrow \alpha_{\text{Tot}} = \arcsin\left(\frac{n_1}{n_2}\right). \quad (323)$$

If a layer with a larger refractive index  $n_2$  is surrounded by layers with a smaller one  $n_1$ , it can be employed as an *optical waveguide*. Light propagating at an angle larger than the angle of total reflection is confined in the middle layer. *Optical fibers* are based on this principal, but light is also guided in a layer which is extending laterally and surrounded by layers with a smaller refractive index.

The angle of total reflection limits the maximum angle  $\theta_{\text{max}}$  at which light may enter or exit an optical waveguide (see Fig. 137). The refractive index  $n_0$  outside of the waveguide is smaller than inside, and the refraction angel  $\alpha_t$  is calculated by Snell's law (322). From Fig. 137, it is seen that  $\sin(\alpha_t) = \cos(\alpha_{\text{Tot}})$ . Snell's law is now:

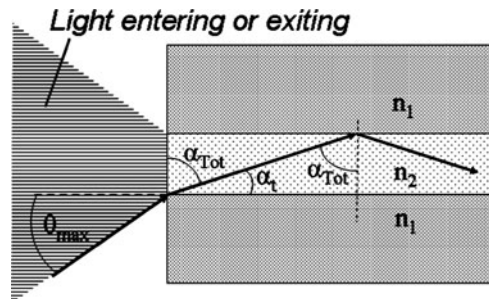
$$n_0 \sin(\theta_{\text{max}}) = n_2 \sin(\alpha_t) = n_2 \cos(\alpha_{\text{Tot}}). \quad (324)$$

The sum of the squares of sine and cosine of every angle  $\alpha$  equals 1:  $\sin^2(\alpha) + \cos^2(\alpha) = 1$ . If the angle of total reflection is inserted into this equation and (323) is used for  $\sin(\alpha_{\text{Tot}})$ , (324) yields:

$$n_0 \sin(\theta_{\text{max}}) = n_2 \cos(\alpha_{\text{Tot}}) = n_2 \sqrt{1 - \frac{n_1^2}{n_2^2}} = \sqrt{n_2^2 - n_1^2} := \text{NA}. \quad (325)$$

The quantity NA defined in the above equation is the *numerical aperture* of the optical waveguide which is a measure of the illumination angle which can be achieved. The numerical aperture is only a function of the refraction indices of the materials it is made of, while the maximum angel  $\theta_{\text{max}}$  at which light may enter into or exit from the waveguide is also a function of the refraction index of the medium outside the fiber.

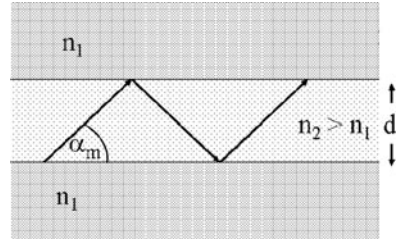
In optical waveguides, the component of the light perpendicular to the mirrors needs to fulfill the condition of standing waves as in a Fabry-Perot interferometer



**Fig. 137** Reflection and refraction of light



**Fig. 138** Propagation of light in a waveguide at mode  $m$



(cf. 321) [78]. As a consequence, the following equation needs to be fulfilled by the light propagating in the waveguide [53]:

$$m \lambda = 2 d \cos(\alpha_m) \Rightarrow \alpha_m = \arccos\left(\frac{m \lambda}{2 d}\right). \tag{326}$$

In the above equation,  $\alpha_m$  is the angle of incidence inside of the optical fiber (Fig. 138). Obviously, light can propagate inside an optical fiber only at distinct angles  $\alpha_m$ . The light propagating at a certain angle is called a mode. The number of modes is limited because the angle of incidence cannot become larger than the angle of total reflection:

$$m \leq \frac{2 d}{\lambda} \cos(\alpha_{Tot}). \tag{327}$$

If the height  $d$  of the waveguide is small, only a small number or even only one mode can propagate inside. The kernel of one mode waveguides typically is approximately  $1 \mu\text{m}$  high, while multimode waveguides have a kernel,  $100 \mu\text{m}$  or more in height. An advantage of one mode waveguides is that all the light is traveling the same distance when propagating through the guide. Light which is reflected at larger angles  $\alpha_m$  of incidence has a longer way to travel, and, therefore, arrives a bit later than the light of other modes. As a consequence, a data bit transmitted through several modes becomes wider, while propagating through an optical fiber. This effect limits the data rate which can be sent through the fiber. Therefore, one mode fibers allow larger data rates.

Equation (327) can be interpreted also in another way: There is a lower limit for the wavelength  $\lambda$  of light which can propagate through a thin waveguide. Thus, a waveguide works as a filter:

$$\lambda \leq \frac{2 d}{m} \cos(\alpha_{Tot}). \tag{328}$$

The light confined in an optical waveguide is not able to leave the middle layer according to classical theory. However, quantum mechanics is needed to describe all aspects of light propagation in waveguides. According to quantum mechanics,

the wave which is employed to describe the light is a measure of the probability that a light quantum, a so-called photon, is found at a certain position. The wave function is not zero in the layer with the larger refractive index but extends into this layer with decreasing amplitude. The part of the wave function which is in the “forbidden” region is called the *evanescent field* (cf. Fig. 139). Thus, there is a little probability that photons exist in the layers with the larger refractive index and even outside of the entire waveguide. As a consequence, the light inside of the waveguide can be affected by media outside.

Due to the evanescent field, there is a way for the light in a waveguide to “tunnel” through the barrier of the lower refractive index into a neighboring region with higher refractive index. After some time and propagation length, the entire wave function has shifted to the other waveguide (cf. Fig. 139a), and after the same distance of propagation it has shifted back into the initial waveguide. A part of the light is absorbed during tunneling, and, therefore, the amplitude of the wave function is reducing.

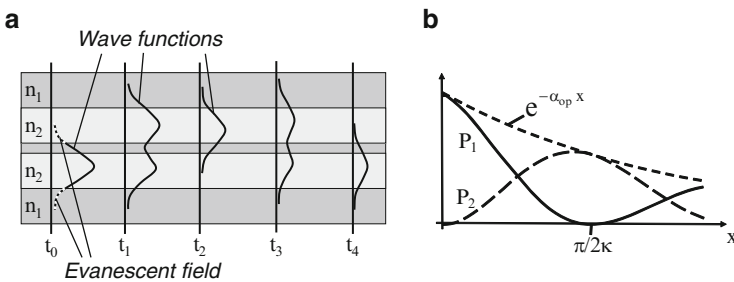
This effect can be employed to distribute light from a single into two waveguides. The power as a function of the position  $x$  in waveguides 1 and 2, respectively, can be calculated with the following equations:

$$P_1(x) = \cos^2(\kappa x)e^{-\alpha_{op}x}, \tag{329}$$

$$P_2(x) = \sin^2(\kappa x)e^{-\alpha_{op}x}. \tag{330}$$

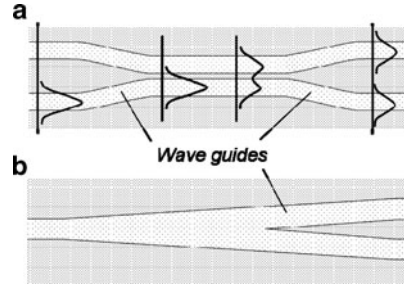
The above equations are plotted in Fig. 139b. The coupling constant  $\kappa$  and the damping constant  $\alpha_{op}$  are a function of the geometry of the waveguides and the modes of the light propagating in them.

When light is to be distributed from one waveguide into two, they need to be designed in parallel for the distance  $\pi/4\kappa$ , as seen from Fig. 139b and (329) and (330). Figure 140 shows such a design and as an alternative a so-called Y-coupler. The Y-coupler is easier to design, because nothing needs to be known about the wave function and modes in the waveguide, but the angle at the junction may not be



**Fig. 139** Light tunneling from a waveguide to a neighboring one (a) and power in the two waveguides (b)

**Fig. 140** Distribution of light from one into two waveguides by (a) tunneling into a neighboring waveguide and (b) a Y-coupler



larger than about  $0.5^\circ$  to avoid too much loss of the light where the required angle of incidence at the change in refraction index cannot be matched.

In some materials, the refractive index  $n_{op}$  is a function of the electric field  $E_{el}$  applied to them. This is called the *electro-optical effect*. By employing this effect, it is possible to modulate the amplitude of light and to switch it on and off. In general, the refractive index can be described by:

$$n_{op} = n_2 - \frac{1}{2} r_P n_2^3 E_{el} - \frac{1}{2} r_K n_2^3 E_{el}^2. \tag{331}$$

In the above equation,  $n_2$  is the refractive index without any electric field and  $r_P$  and  $r_K$  are material constants: the Pockels constant and the Kerr constant, respectively. Two kinds of the electro-optical effect are distinguished: the *Pockels effect* and the *Kerr effect*. The Pockels effect is linear with the electrical field. It is represented by the term in (331), which is linear with the electric field. The other electro-optical effect is the Kerr effect. With the Kerr effect, the refractive index of a material is changed with the square of the electric field [the last term in (331)]. The part of an electro-optical material in which the Pockels or the Kerr effect is employed is called a Pockels cell or a Kerr cell, respectively.

Tables 17 and 18 show some materials with comparatively large Pockels and Kerr constants, respectively.

As described above, the refractive index is the ratio of the velocity of light in a certain material and in vacuum. Therefore, decreasing the refractive index results in a smaller delay of the propagating light.

This is employed in a device called *Mach-Zehnder interferometer* (cf. Fig. 141). Light coming from the left is distributed into two waveguides and changed in one of them. After reunification of the waveguides, the two waves interfere resulting in a changed amplitude.

If a Pockels cell is installed in one of the arms of a Mach-Zehnder interferometer, the delay of the wave in that arm can be manipulated by an electric signal and, this way, the amplitude of the outgoing light is controlled. The phase shift  $\Delta\varphi$  in a Pockels cell with length  $L_{PZ}$  of a wave which shows a wavelength of  $\lambda_0$  in vacuum is according to (331):

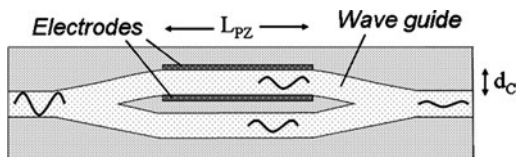
**Table 17** Pockels constants

Material with Pockels effect	Pockels constant $r_p$
Potassium dihydrogen phosphate ( $\text{KH}_2\text{PO}_4$ )	$3.6 \times 10^{-11} \text{ m/V}$
Deuterated potassium dihydrogen phosphate ( $\text{KD}_2\text{PO}_4$ )	$8 \times 10^{-11} \text{ m/V}$
Lithium niobate ( $\text{LiNbO}_3$ )	$37 \times 10^{-11} \text{ m/V}$

**Table 18** Kerr constants

Material with Kerr effect	Kerr constant $r_K$
Nitrobenzene	$2.4 \times 10^{-12} \text{ m/V}^2$
Glasses	$3 \times 10^{-16} \text{ to } 2 \times 10^{-25} \text{ m/V}^2$
Water	$4.4 \times 10^{-14} \text{ m/V}^2$

**Fig. 141** Mach-Zehnder interferometer with a Pockels cell (region between the electrodes) in one of its arms



$$\Delta\varphi = \frac{\pi L_{PZ}}{\lambda_0} r_p n_2^3 E_{el}. \quad (332)$$

The light is switched off when the phase shift of the wave is  $\Delta\varphi = \pi$ . Therefore, the above equation yields for switching off the light after a Mach-Zehnder interferometer with a Pockels cell:

$$E_{el} = \frac{\lambda_0}{L_{PZ} r_p n_2^3} \Rightarrow U = \frac{d_C \lambda_0}{L_{PZ} r_p n_2^3}. \quad (333)$$

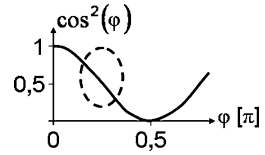
In the above equation,  $U$  and  $d_C$  are the voltage applied to the electrodes and their distance, respectively. Light switching with a Pockels cell is very quick. Thus, light pulses, only a few nanoseconds in length, can be generated and sent as data bits into an optical fiber.

It is also possible to modulate the intensity of the light by changing the phase shift in a Pockels cell. If  $I_e$  and  $I_a$  denote the intensity of the light at the input and output of the cell, respectively, the modulation is described by:

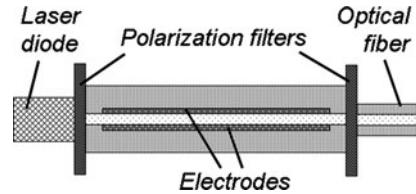
$$I_a = I_e \cos^2\left(\frac{\pi L_{PZ}}{\lambda_0} r_p n_2^3 U\right). \quad (334)$$

The above equation is shown in Fig. 142. When light is to be modulated with a Pockels cell, the nearly linear part of the  $\cos^2$ -function around  $\pi/4$  is used.

**Fig. 142** Linear part of the  $\cos^2$ -function marked by a dashed ellipse



**Fig. 143** Device employing turning of polarization direction of light for sending data bits into an optical fiber



There are also devices employing the effect that some materials turn the polarization angle of light (cf. Fig. 143). Before entering into the device, the light passes a polarization filter. At the end of the device, a second polarization filter is arranged adjusted for light polarized perpendicular to the first filter. Thus, only light with a polarization direction turned by  $90^\circ$  in the cell can pass it. This is also a way to generate data bits which are sent through an optical fiber.

## Exercises

### Problem 27

The resolution of an optical instrument is defined as the distance between neighboring points of the image which can be distinguished from each other. According to Lord Rayleigh, two points can be distinguished from each other, if the intensity maximum of one point lies in the minimum of the neighboring one.

- (a) Find a way to calculate the resolution of an ideal lens on the basis of (314) (page 177). Derive an equation for the calculation of the resolution limit.
- (b) It does not help to make the distance of the pixels of a CCD chip narrower than the resolution of the lens. The lens of the digital camera of a mobile phone has a diameter of 2 mm and a focus length of 28 mm. Calculate for the wavelength of visible light with the maximum intensity in the sun light (550 nm) the minimum distance of pixels which makes sense.
- (c) For comparison, calculate the resolution of the lens of a 35 mm camera at the same wavelength. Assume that the lens of that camera has a diameter of 30 mm and a focal length of 50 mm.

# Diffusion

If two liquids are to be mixed in macroscopic applications, it is usual to generate some turbulence which facilitates mixing very much. Turbulence is achieved when the *Reynolds' number*  $Re$  becomes larger than approximately 1,500. The Reynolds' number is the ratio of inertial and friction forces in a flow. The inertial forces are described by the product of the density  $\rho_F$  of the liquid, a characteristic length  $L$  of the vessel or a rigid structure in interaction with the flow, and the velocity  $v$  of the flow. The frictional forces are described by the viscosity  $\eta$  of the fluid:

$$Re := \frac{\rho_F L}{\eta} v. \quad (335)$$

As an example, the Reynolds' number of a flow in a water pipe can be estimated: If the flow is generated by a pressure difference of 50 kPa over a pipe, 1,000 m in length, and 1 cm in diameter, with (177) (on page 108), and the viscosity of water ( $10^{-3}$  Pa s), the mean velocity of the water is calculated to be 16 cm/s. By inserting this into (335) together with the density of water ( $1,000 \text{ kg/m}^3$ ), the Reynolds' number of 1,562 is obtained. This result clearly shows that the flow in a water pipe is at the limit of a turbulent flow. If we reduce all dimensions by a factor of 1,000, the diameter, length, and mean flow velocity become  $10 \text{ }\mu\text{m}$ , 1 m, and  $160 \text{ }\mu\text{m/s}$ , respectively. As a result, Reynolds' number is 0.0016. This example shows that in most cases there is no turbulent flow in microcapillaries. As a consequence, mixing of fluids in microtechnique can only hardly be achieved by turbulent flow. The only remaining possibility is micromixing by diffusion. Fortunately, mixing becomes quicker when dimensions are scaled down.

Diffusion is due to the motion of atomic and molecular particles. The temperature of all substances is due to the average kinetic energy of its atomic and molecular particles. This kinetic energy is the consequence of a random movement of the particles. That is, the higher the temperature  $T$  of a substance is, the quicker are its particles moving:

$$\frac{1}{2} m_T \langle v^2 \rangle = \frac{3}{2} k_B T \Rightarrow \langle v^2 \rangle = 3 \frac{k_B T}{m_T}. \quad (336)$$

In the above equation, the left most term is the average kinetic energy of the particles. The angle bracket indicates that the average of the square of the velocities has to be calculated.  $k_B$  denotes the Boltzmann constant ( $=1.4 \times 10^{-23}$  J/K).

Since the molecular particles are moving randomly, the average of the velocity and the average of the moved distance in a certain time are both zero. Only the average of velocity and moved distance are not zero. Einstein was the first who calculated the average of the square of the moved distance  $s$  during a time  $\Delta t$  of particles with radius  $R_T$  in a fluid with viscosity  $\eta$ :

$$\langle s^2 \rangle = \frac{k_B T}{3 \pi \eta R_T}. \quad (337)$$

The random walk of small particles in a liquid can be observed with a microscope. It is called Brownian motion.

Brownian motion of atomic and molecular particles is the reason for diffusion. If two fluids (gases or liquids) are in contact to each other, the random walk of their particles results in mixing of the two fluids. Diffusion of gases and liquids also occurs in rigid bodies when, e.g., the particles of a gas move between the atoms of the crystal lattice.

Diffusion can be calculated by solving the following differential equation which is Fick's second law:

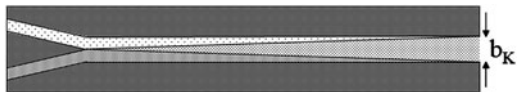
$$\frac{\partial c_F}{\partial t} = -D \frac{\partial^2 c_F}{\partial x^2} \quad \text{with } D = \frac{k_B T}{6 \pi \eta R_T}. \quad (338)$$

In the above equation,  $D$  is the diffusion constant as calculated by Einstein from the random walk of particles with radius  $R_T$  moving in a fluid with viscosity  $\eta$  at temperature  $T$ .

Besides the fact that the mean of the velocity squares of the molecular particles is rising with temperature, (336) shows that smaller particles (with less mass) are moving faster at the same temperature. As a consequence, very small movable parts tend to move randomly generating a random noise, e.g., of an acceleration sensor such as shown in Fig. E15 on page 83. Therefore, there is a limit for reducing the size of such sensors.

If two fluids with the same density and wettability (capillary forces are the same) are fed into a single microcapillary, they follow a laminar flow next to each other (cf. Fig. 144). While they are in contact with each other, diffusion starts and downstream they get increasingly mixed. There is no equation with which the time  $\Delta t$  for mixing in a capillary with width  $b_K$  in a medium with diffusion constant

**Fig. 144** Mixing two fluids coming from the left in a microcapillary



D can be calculated exactly. However, as a rule of the thumb the following equation is valid [54]:

$$\Delta t \approx \frac{b_K^2}{4 D}. \tag{339}$$

Often it is desirable to reduce mixing time and the overall dimensions of the mixing capillary, because both result in higher costs. Reducing the flow velocity requires a shorter capillary; however, the amount of mixed fluid per time is diminished that way. A large mixing rate is wanted, and, therefore, the objective is to reduce mixing time and overall dimensions at a given flow rate and cross-sectional area of the capillary.

A possibility to increase mixing is to raise temperature. According to (338), the diffusion constant is proportional to temperature. Besides this, the viscosity of most fluids decreases very much with rising temperature resulting in a diffusion constant which is increasing nearly with the square of temperature. However, enhancing temperature is often limited because liquids should not boil and, e.g., organic substances can be destroyed at comparatively low temperatures.

Another possibility is to keep the cross-section or hydraulic diameter [cf. (176) on page 107] of the mixing capillary constant by decreasing the width  $b_K$  (the direction of diffusion) and increasing the height. This way, the contact area of the fluids to be mixed is increased and the necessary diffusion depth is reduced. Reducing the width of a microcapillary is limited because the risk that particles block it, is higher for narrower channels. Often a minimum width of 100  $\mu\text{m}$  is a good compromise. Since it is desirable to reduce the mixing time even more than what can be achieved with a reduced width at constant cross-section, a lot of designs have been developed which all enhance the contact area and reduce the necessary diffusion depth. In the following, some of these designs are introduced.

Instead of feeding the fluids into a single capillary in parallel, they can be fed in sequentially (cf. Fig. 145a). Due to the parabolic profile of the laminar flow, the contact area of the fluids is increasing downstream rapidly and the necessary diffusion depth is drastically reduced. If the width of a parabola at the tip of the following one is taken as an approximate measure for the necessary diffusion depth  $d$ , an equation of the reduction of mixing time is obtained. The length  $x$  of the

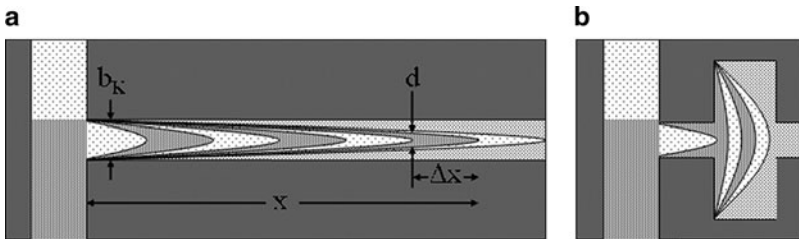


Fig. 145 Sequential mixing without (a) and with (b) increase of capillary diameter



parabola from the beginning of the capillary to the peak is calculated from a parabola equation with the unknown parameter  $a$ :

$$x = a \left( \frac{b_K}{2} \right)^2 \Rightarrow a = 4 \frac{x}{b_K^2}. \quad (340)$$

The distance  $\Delta x$  from the peak of one parabola to the next one is according to the above parabola equation:

$$\Delta x = a \frac{d^2}{4} = \frac{x}{b_K^2} d^2 \Rightarrow \frac{d^2}{b_K^2} = \frac{\Delta x}{x} = \frac{1}{n}. \quad (341)$$

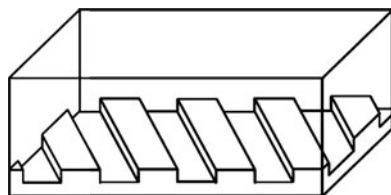
In the above equation,  $n$  is the number of switchings between the two fluids during sequential feeding. The comparison with (339) shows that the mixing time reduces approximately proportional to the inverse of the number of switchings.

Another solution for reducing mixing time is enhancing the cross-sectional area of a channel with sequential feeding (cf. Fig. 145b). Due to the laminar flow, the volume of every sequence is spread over the larger diameter and diffusion is enhanced. If the diameter is increased by a factor of 10, the cross-section is enhanced and the diffusion depth is reduced by a factor 100, and the mixing time according to (339) is shortened by a factor 10,000. That is, the mixing time scales with the inverse of the fourth power of the diameter change.

The disadvantage of micromixers with sequential feed is that some means is required which provides the switching between the two fluids.

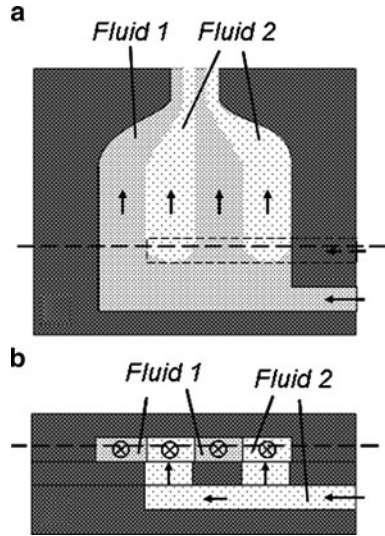
Mixing with parallel feed can be enhanced by ridges on at least one of the walls of the mixing channel (see Fig. 146) [55]. The ridges are designed at an angle (e.g., 45°) to the flow direction, and, this way, generate a rotation in the laminar flow increasing the contact area between the two fluids. As a result, diffusion and mixing are enhanced significantly. The fabrication of ridges on one wall of a channel can be comparatively easy, e.g., by polymer molding. Unfortunately, there is no equation available with which the mixing time can be calculated easily. Thus, Finite Element calculations or experiments with varying dimensions are necessary to find the optimum design.

Mixer designs which are easier to understand are those which separate the parallel flow of two incoming fluids and join them again in a changed arrangement such that the interface area between the fluids is enhanced. Figure 147 shows a possible



**Fig. 146** Section of a mixing channel with ridges on one wall

**Fig. 147** Micromixer injecting one fluid into the flow of the other one



design. Fluid 1 is entering a wide channel (cf. 147a). From a lower level, Fluid 2 is entering through several (In Fig. 147 there are shown two.) orifices into the wide channel. This way, several flow stripes are generated next to each other. These stripes are then entering a narrower channel which brings them into close contact. Except the stripes directly at the wall, diffusion now can start to both sides of the stripe. This halves the necessary diffusion depth, and according to (339), the approximate mixing time as a function of the number of stripes  $n$  generated becomes:

$$\Delta t \approx \frac{4 b_K^2}{4 n^2 D} = \frac{b_K^2}{n^2 D}. \tag{342}$$

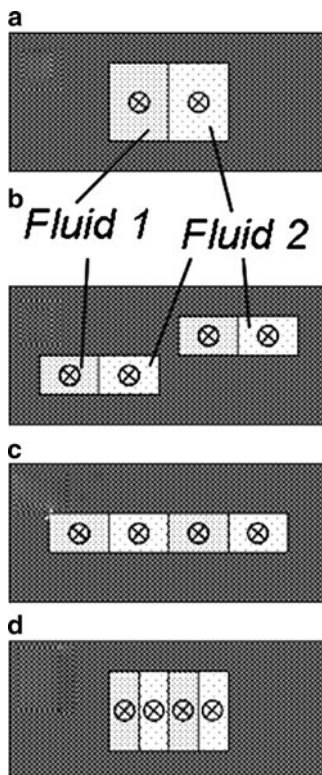
The above equation results in a reduction of the mixing time by a factor of  $4/n^2$  compared with mixing in a channel without generating several stripes.

Owing to the parabolic flow profile, it may be an advantage to design the flow stripes in the center of the wide channel narrower than at the wall, because the quicker flow results in less time for diffusion.

The width of the wide channel is determined by the minimum width of the orifices between the channel of fluid 2 and the wide channel. This width is required to avoid blocking by particles which may occur in the fluid. When the narrowing of the wide channel to the same width is designed, it needs to be taken into account that bubbles can block a channel where it gets narrower (cf. page 118f).

Another way to generate flow stripes is to separate a channel with the parallel flow of two liquids vertically into two flow channels containing the parallel flow of both fluids, Figure 148a, b shows cross-sections of this channel. Then the two

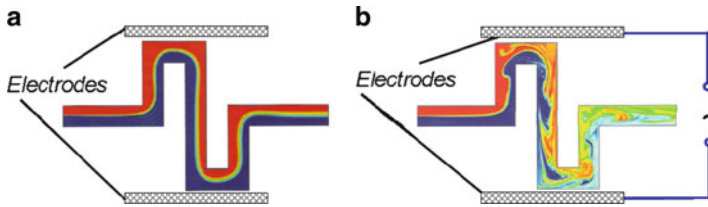
**Fig. 148** Vertical separation and horizontal joining a flow



channels are joined horizontally (Fig. 148c) and the width is reduced to the half and the height is doubled achieving the dimensions of the initial channel again (cf. Fig. 148d). This procedure can be employed several times, each time reducing the diffusion depth by a factor of 2. If, as an approximation, it is assumed that mixing starts not before the last step of dividing and reunifying the flow, the mixing time can be calculated with (342), because a number of parallel flow stripes is generated there also.

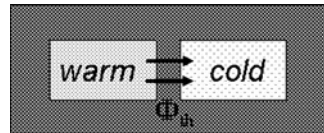
So far passive micromixers have been described. Mixing can be supported also by active devices which generate an enlarged interface area between the fluids to be mixed. For example, ultrasound can be employed in a mixing chamber or an electro-osmotic motion (cf. page 122) can be overlaid to a flow around corners or other structures in the channel [56]. Figure 149 shows the FEM calculation of the flow of two liquids through a curved channel. Owing to the laminar flow, only slow mixing by diffusion occurs (a). If, however, an alternating electric field is applied to the channel at the region of the curves, fast mixing is observed (b).

Not only atoms and molecules are exchanged by diffusion but also the temperature itself. As heat conduction is a diffusion process, the concepts employed for mixing also help to enhance heat exchange. The interface area between the media



**Fig. 149** FEM calculation of slow mixing in a curved flow channel (a) and fast mixing in the same channel (b) with an alternating electro-osmotic motion overlaid to the flow [56]. (Courtesy of Karlsruhe Institute of Technology, KIT)

**Fig. 150** Heat flow through the wall between channels



between which heat is to be exchanged needs to be designed as large as possible. If a fluid is used to cool or heat another fluid but both fluids shall not be mixed, flow stripes need to be separated by walls which are as thin as possible. Thus, parallel channels have to be constructed which are very near to each other in a material which is a good heat conductor. Such devices are called *heat exchangers*.

The heat flow  $\Phi_{th}$  from a warm channel to a cold one with a temperature difference  $\Delta T$  through a wall with area  $A$  (cf. Fig. 150) is described by the following equation:

$$\Phi_{th} = k_w A \Delta T. \tag{343}$$

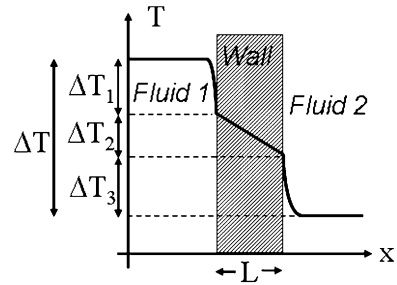
The quantity  $k_w$  in the above equation is the heat transition coefficient. The heat flow per volume of the heat exchanger is obtained by dividing (343) by the volume  $V$ :

$$\frac{\Phi_{th}}{V} = k_w \frac{A}{V} \Delta T. \tag{344}$$

The ratio of the area of the walls in a microreactor to its volume obviously becomes larger when the overall size is reduced (cf. page 3), providing a much larger heat exchange than in a macroscopic heat exchanger. Besides this, the heat transition coefficient also increases with decreasing dimensions.

The heat transition through a wall consists of three steps (Fig. 151): The heat transfer from the warmer fluid to the wall, the heat conduction through the wall, and the heat transfer to the colder fluid. When the temperatures of the fluids are held constant, the heat flow for all three steps is the same resulting in the following equation:

**Fig. 151** Temperature distribution due to heat flow through a wall



$$\Phi_{\text{th}} = \alpha_1 A \Delta T_1 = \frac{\lambda_{\text{th}}}{L} A \Delta T_2 = \alpha_2 A \Delta T_3. \quad (345)$$

In the above equation,  $\lambda_{\text{th}}$ ,  $L$ ,  $\alpha_1$ ,  $\alpha_2$ , and  $\Delta T_i$  are the heat conductivity of the wall, the thickness of the wall, the heat transfer coefficients of fluid 1 to the wall and from the wall to fluid 2, and the temperature differences, respectively, as shown in Fig. 151. The sum of the temperature steps is the total temperature difference between the fluids:

$$\Delta T = \Delta T_1 + \Delta T_2 + \Delta T_3. \quad (346)$$

Equations (345) and (346) result in the following expression:

$$\Delta T = \frac{\Phi_{\text{th}}}{\alpha_1 A} + \frac{L \Phi_{\text{th}}}{\lambda_{\text{th}} A} + \frac{\Phi_{\text{th}}}{\alpha_2 A} = \frac{\Phi_{\text{th}}}{A} \left( \frac{1}{\alpha_1} + \frac{L}{\lambda_{\text{th}}} + \frac{1}{\alpha_2} \right) = \frac{\Phi_{\text{th}}}{A} \frac{1}{k_{\text{W}}}. \quad (347)$$

The last term in the above equation is obtained from the comparison with (343) yielding:

$$\frac{1}{k_{\text{W}}} = \frac{1}{\alpha_1} + \frac{L}{\lambda_{\text{th}}} + \frac{1}{\alpha_2}. \quad (348)$$

The heat transfer coefficients  $\alpha_i$  can be expressed by the heat conductivity  $\lambda_{\text{F}}$  of the fluid, the Nusselt number  $\text{Nu}$ , and the hydraulic diameter  $D_{\text{h}}$  of the flow channel:

$$\alpha_i = \frac{\text{Nu} \lambda_{\text{F}}}{D_{\text{h}}}. \quad (349)$$

The Nusselt number is a parameter which describes the heat transfer between a flowing fluid and a rigid surface. It is a function of the flow velocity and material constants. Therefore, from (348), it is concluded that decreasing the size of channels, and, that way reducing the hydraulic diameter and the width of the walls, results in an increase of the heat transition coefficient  $k_{\text{W}}$ . The increase in heat flow

per volume as calculated with (344) due to the larger heat transition coefficient and the surface to volume ratio can become much larger than a factor 100.

Heat exchangers can also be employed to control the temperature very precisely at which a chemical reaction takes place. In a microreactor mixing channels can be surrounded by channels through which a fluid (e.g., water) at the desired temperature is flowing comparatively quickly. This way, even exothermal reactions which under other conditions would result in an explosion can be held under control, such as the reaction of oxygen and hydrogen to water.

Although microreactors have comparatively small outer dimensions, their throughput can be large and suitable for large scale industrial production, because the reaction time is nearly the same as the mixing time which according to (339) (page 193) is proportional to the square of the channel width. The channel width can be made much smaller in microtechnique than in macroscopic reactors. Besides this, the chemical reaction at well-controlled temperature yields the desired product with high efficiency.

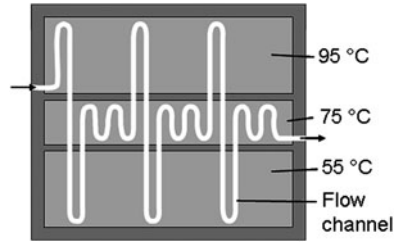
A special type of microreactor is a chip for polymerase chain reaction (PCR). PCR is the biological reaction which is employed to enhance the number of DNA molecules to ease their analysis. Such PCR chips are under development at different laboratories. If once they are available at low cost, they could revolutionize medical diagnosis. Instead of analyzing the symptoms of a patient, it could become possible to detect the DNA of the virus or bacteria in the blood. Certainly, other applications would benefit from a generally available and low-cost DNA analysis also, such as forensic analysis, paternity testing, and detection of hereditary diseases.

For PCR, the helix of the DNA needs to be split into two single strands. This is achieved by heating the sample up to 95°C for 1 min. Then a special kind of molecules, so-called primers, which had been added to the sample, needs to be affixed to certain positions of the DNA. This happens when the sample is held at 55°C for 45 s. In a third step nucleotides, the basic acids from which the DNA is build up (adenine, guanine, cytosine, and thymine) starting at the primers are added according to the corresponding acids on the existing DNA strand to complete it to a double helix again. This is achieved when the temperature of the sample is held at 75°C for 2 min. This way, in a thermal cycle, one double helix is transferred into two. When the thermal cycle is performed several times, e.g., 20 or 40 times, from one DNA there are generated millions of copies which in following steps can be analyzed.

As a consequence, a PCR chip needs to thermally cycle the sample between three temperatures needed for the process. A simple way to achieve this is to let the sample flow through a channel which passes through temperature zones (cf. Fig. 152). The length of the channel in the temperature zones is adapted such that the sample stays there for the required time. It would also be possible to change the cross-section of the channel in the different zones to adapt the flow velocity. However, a changing cross-section enhances the risk that the channel is blocked by particles in the liquid or by bubbles (cf. page 118f).

In general, the largest problem of microheat exchangers and microreactors is blocking of the microchannels by particles, algae growing inside, and bubbles.

**Fig. 152** Schematic drawing of a PCR chip with only three thermal cycles and three temperature zones



Besides this, corrosion may be a problem because the large area of the thin walls is quickly corroded away.

### Exercises

#### *Problem 28*

Two different water-based solutions are fed into a mixing channel. The flow velocity in the channel is 1 cm/s. The radius of the diluted molecules is 0.1 nm and the width of the mixing channel is 100 μm.

- (a) How long must the mixing channel be designed to achieve complete mixture?
- (b) For comparison consider a water pipe into which a colorant is leaking. The flow velocity and the radius of the colorant molecules is the same as in Question (a). The diameter of the water pipe is 1 cm. After what flow distance are water and colorant completely mixed?

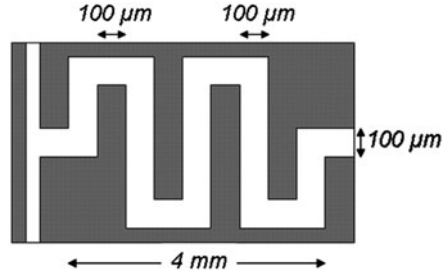
Hint: Use the diameter of the water pipe as the approximate diffusion depth in the equation.

Dynamic viscosity	1.002 mPa s	Boltzmann constant	$1.38 \times 10^{-23}$ J/K
Temperature	298 K		

#### *Problem 29*

With the micromixer shown in Fig. 144 (page 192), ethanol and water are mixed in a 1:1 ratio. The flow velocity of both liquids together is 20 μL/min. The cross-section of the mixing channel is 100 μm × 100 μm. The diffusion constant of ethanol in water is at room temperature  $0.84 \times 10^{-5}$  cm<sup>2</sup>/s. The mixing channel is arranged such that the lighter ethanol is flowing above the water. This way gravity nearly does not influence the mixing process. (Otherwise the ethanol would go to the upper side first and during this mixing can be facilitated.)

**Fig. E24** Winding mixing channel



- (a) Please calculate how long the mixing channel needs to be to obtain complete mixture of the two liquids.
- (b) The mixer from (a) shall not be longer than 4 mm except the inlet and the outlet part. It shall be arranged in windings on an area which has to be as small as possible and the distance between the channels cannot be less than 100 μm because of the fabrication process (see Fig. E24). Please calculate how many windings are necessary and how wide the mixer needs to be. Ignore the area needed around the mixer.
- (c) Another possibility to reduce the size of the mixer is to design a narrower channel. Please calculate how wide a mixing capillary needs to be to achieve complete mixing of ethanol and water in a 4-mm long straight channel. The height of the channel remains to be 100 μm.
- (d) Assume now that the height of the capillary in (c) is increased to 2 mm. How long must this channel be if its width is 4 μm?
- (e) The capillary with a cross-section of 100 μm × 100 μm shall now be split into several parallel flow stripes and reunified as shown in Fig. 147 on page 195. The other conditions are not changed. How many parallel flow stripes are necessary to achieve complete mixing in a 200-μm long channel?
- (f) Approximately the same mixing rate as in (e) is achieved, if the flow is split vertically into two equal parts and reunified horizontally as shown in Fig. 148 so many times as there are parallel flow stripes in (e). Make a justified guess which of the two possibilities needs more space, and, therefore, is more expensive. Assume for simplicity that the mixing starts not before the last reunification of the flow paths.



# Microvalves

Microvalves have a lot of potential applications. They can be employed for dosing of small liquid volumes, controlling pneumatically or hydraulically driven robots or machines, and as a pilot valve which switches a larger valve.

Potential advantages of microvalves compared with macroscopic valves are a smaller size, less weight, less cost, smaller dead volume, shorter switching time, and less energy consumption. The main problems which occur with microvalves are small flow through and large pressure drop over the open valve, blocking by particles and bubbles, and sensitivity to temperature changes. The reason for these problems is mostly that it is difficult to build an actuator which provides both large stroke and large force, i.e., the energy generated in a small volume is too small.

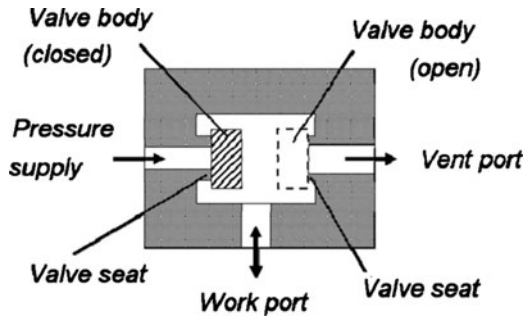
Valves, in general, are classified by the number of ports and the number of switching states they have. In Fig. 153, a 3/2-way valve is shown. This means that there are three ports and two possible switching states which correspond to two possible positions of the valve body. If the valve body is pressed onto the valve seat of the port which is connected to the pressure supply (on the left in Fig. 153), the valve is closed and no medium can go into the work port. In this state, the work port can be vent through the valve which may be necessary, e.g., to move a piston back into its rest position.

If the valve body is pressed onto the seat of the vent port, the valve is open and the pressure from the supply reaches the work port, and e.g. let a piston make a stroke. The vent port is closed in this switching state of the valve.

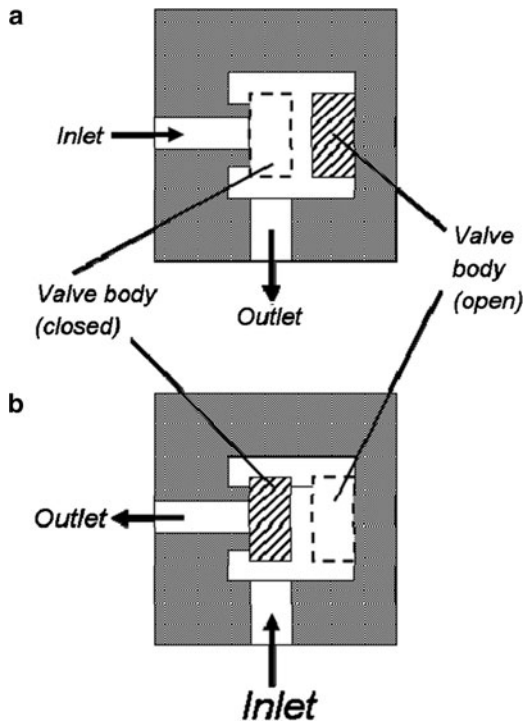
3/2-way valves are a very common design for valves which are employed to control the motion of devices. More basic is the design of 2/2-way valves which only open or close a flow or pressure supply. A 3/2-way valve can be composed of two 2/2-way valves when the inlet of the one valve and the outlet of the other valve are connected to the work port and one valve is opened when the other one is closed and vice versa. There are also more complicated designs such as 6/2-way valves, but they are needed only for special applications and all can be replaced by a suitable combination of 2/2-way valves.

The principal design of a 2/2-way valve is shown in Fig. 154. The two designs shown differ only in the direction of the flow. As a consequence, the valve in Fig. 154a is open, when the actuator is not powered and the other design results in

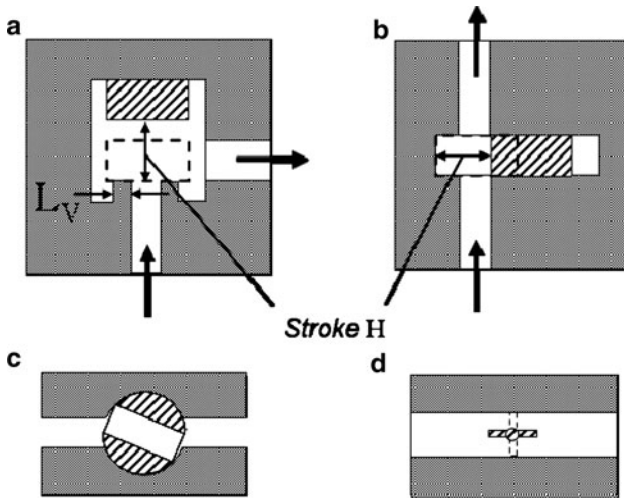
**Fig. 153** Schematic drawing of a 3/2-way valve. The valve body is either in the closed or the open position



**Fig. 154** Schematic drawing of a 2/2-way valve as normally open (a) and normally closed type (b)



a closed valve when there is no power. These two types of valves are called normally open and normally closed, respectively. There are macroscopic valves which are normally open or normally closed independent on which port the higher pressure is applied. There are applications for which normally closed valves are needed such as the valves at a pressure or gas support which needs to close to avoid losses when there is a power break down. For other applications, normally open valves are required, e.g., a vacuum vessel should be vent for safety reasons when power is not available. Besides normally open and normally closed valves, there are also bistable valves which retain their state when power is turned off.



**Fig. 155** Valve types where the valve body is moving in flow direction (a), perpendicular to it (b), or turned in the flow as a channel (c) or a flap (d)

Valves are also distinguished from each other by their working principle. The valve body may be moved in flow direction (seat valve, cf. Fig. 155a), perpendicular to the flow (slide valve, cf. Fig. 155b), or it can be turned in the flow either as a flap (Fig. 155d) or as a plug (Fig. 155c). The latter design is not suitable for microtechnique because the sealing between two rigid walls can hardly be achieved and the friction between them is comparatively large and hard to overcome by a microdrive.

The fabrication of a flap appears to be possible, but it is difficult to connect a suitable drive to it. A drive inside of the medium is not desirable, because the medium could affect the drive or the drive could heat the medium or influence it in a different way. If the drive is placed outside of the flow channel, it is difficult to seal the mechanical connection to it. Besides this, a 90° turn is a comparatively large stroke for a microactuator. Therefore, in the following only the valve types shown in Fig. 155a, b are discussed in more detail.

If a seat valve is to be closed against a certain pressure difference  $\Delta p$ , the force  $F_H$  needs to be overcome by the actuator which equals the product of the pressure difference and the cross-sectional area  $A$  of inner edge of the valve seat. To close the valve, the work  $W_H$  needs to be done by the actuator which is calculated from the product of the force  $F_H$  and the stroke  $H$  of the valve:

$$F_H = \Delta p A \quad \text{and} \quad W_H = F_H H = \Delta p A H. \tag{350}$$

The narrowest part of the flow channel through the valve is the gap between valve body and valve seat. The flow through the valve is controlled by the height of

this gap. If a laminar flow is assumed, the pressure drop  $\Delta p_s$  over the valve seat can be calculated with the Hagen Poiseuille equation [(175) on page 107]:

$$\Delta p_s = -32 \frac{\eta L_V}{D_h^2} v \Rightarrow v = \frac{-D_h^2}{32 \eta L_V} \Delta p_s. \quad (351)$$

As an approximation, it is assumed here that the gap of the valve corresponds to a capillary with a length  $L_V$  which is equal to the width of the valve seat and a width which is equal to the circumference length  $U_S$  of the valve gap. As a consequence, the hydraulic diameter  $D_h$  [cf. (176) on page 107] is calculated as:

$$D_h = \frac{4 H U_S}{2 U_S} = 2 H. \quad (352)$$

The volume flow  $\Phi_F$  through the valve is calculated from the product of the mean velocity  $v$  of the flow and the cross-sectional area  $A$  of the gap. With (351) and (352) now it is obtained:

$$\Phi_F = v A = v H U_S = \frac{-D_h^2 H U_S}{32 \eta L_V} \Delta p_s = \frac{-H^3 U_S}{8 \eta L_V} \Delta p_s. \quad (353)$$

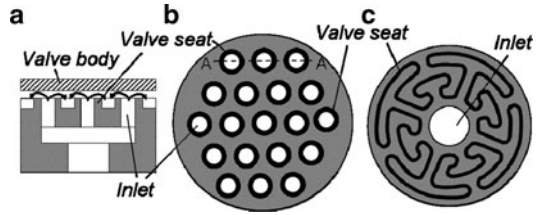
It needs to be noted that in the above calculation, it is ignored that the direction of flow is often changed inside of a valve. Besides this, it is assumed above and in the following that the flow would be laminar. However, especially the flow of a gas can be turbulent in comparatively wide and short channels of a microvalve. As a consequence, the real flow can be much smaller (e.g. a factor of 10) than expected from (353).

Although the real flow can be much smaller than calculated with the equations presented here, the principle interrelationships are shown the right way, and, therefore, help to understand how microvalves should be designed.

A microdrive typically is not able to provide a large stroke and a small stroke results in quick switching of the valve, because the valve body needs to be moved only a small distance. The above equation shows that the stroke can be reduced when the circumference length of the seat is enhanced and its width is reduced.

The minimum width of the seat of a microvalve in most cases is limited by the minimum dimensions which can be fabricated. Contrary to this, the length of the circumference can be enlarged, e.g., by employing more than one valve seat as shown in Fig. 156a, b. Another possibility is to design a winding valve seat such as shown in Fig. 156c [58]. In both cases, the disadvantage is that the cross-sectional area of the valve seat is increased. Therefore, a larger actuator force is required; and it is also a problem for many microactuators to generate large forces. Thus, it can be concluded that the product of stroke and force which is generated by a suitable actuator needs to be large enough, which according to (350) is the work or energy output of the actuator. Besides this, the stroke  $H$  enters into (353) with the third

**Fig. 156** Valves with an increased length of the circumference of the valve seat. Cross-section (a) and top view of multiple inlets (b), and top view of a winding valve seat (c) [58]



power and, therefore, it is hard to compensate a smaller stroke by changing other dimensions, and it remains to be a challenge for designers to achieve a large enough actuator stroke.

In the above and the following equations, it has not been taken into account that the flow direction is changed along the flow path. As a result, there is a much larger flow resistance than calculated, and it should not be expected that the calculated flow is really achieved. However, the principle interrelationships are described correctly.

For an open valve, the pressure drop over the valve seat is not the same as the one over the entire valve, because some feed channels are required to connect the ports to the valve seat. As an approximation, it is assumed that the feed channels have a constant cross-section  $A_Z$  with hydraulic diameter  $D_{h,Z}$  and an overall length  $L_Z$ . The volume flow through the entire valve  $\Phi_F$  is equal to the flow through the valve gap  $\Phi_S$  and the feed channels  $\Phi_Z$ , and the pressure drop  $\Delta p_V$  over the entire valve is the sum of the pressure differences over the valve seat  $\Delta p_S$  and the feed channels  $\Delta p_Z$ . According to (353) and (351), the following equations are obtained:

$$\begin{aligned} \Delta p_V &= \Delta p_S + \Delta p_Z = \frac{8 \eta L_V}{-H^3 U_S} \Phi_S + \frac{32 \eta L_Z}{-D_{h,z}^2 A_Z} \Phi_Z \\ &= - \left( \frac{L_V}{H^3 U_S} + \frac{4 L_Z}{D_{h,z}^2 A_Z} \right) 8 \eta \Phi_F \end{aligned} \tag{354}$$

$$\Rightarrow \Phi_F = \frac{-\Delta p_V}{8 \eta \left( L_V / (H^3 U_S) + 4 L_Z / (D_{h,z}^2 A_Z) \right)} =: \frac{-\Delta p_V}{R_{fl}}. \tag{355}$$

The comparison of (355) with (353) yields the interrelationship between the pressure drop over the valve seat and the entire valve:

$$\Delta p_S = \frac{1}{1 + \left( 4 H^3 U_S L_Z / (D_{h,z}^2 A_Z L_V) \right)} \Delta p_V. \tag{356}$$

For an annular valve, seat with radius  $R_Z$  and width  $L_V$  this is:

$$\Delta p_s = \frac{R_Z^3}{R_Z^3 + 2 H^3(L_Z/L_V)} \Delta p_v. \tag{357}$$

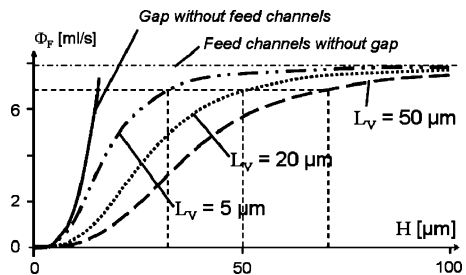
For a square valve, seat with edge length  $H_Z$  and width  $L_V$  this is:

$$\Delta p_s = \frac{H_Z^3}{H_Z^3 + 16 H^3(L_Z/L_V)} \Delta p_v. \tag{358}$$

The quantity  $R_{fl}$  in (355) is called the flow resistance of the valve. It will be used later in this book. The flow as a function of the stroke is called the characteristic curve of a valve. The characteristic curve of a microvalve calculated with (355) is shown in Fig. 157 as a function of the width of the valve seat  $L_V$ . Circular feed channels and a circular valve seat both with a radius of 100  $\mu\text{m}$ , a length of the feed channels of 500  $\mu\text{m}$ , a pressure drop of 100 kPa over the valve, and water as the fluid ( $\eta = 1 \text{ mPa s}$ ) have been assumed.

The flow through the microvalve is increasing with the stroke  $H$  and asymptotically approaching the theoretical limit of a valve without any slit which is obtained from (355) with an infinitely large stroke or an infinitely small width  $L_V$  of the valve seat. As it is difficult to achieve a large stroke, the question arises what minimum stroke is required. For large strokes, a further enhancement does not yield a significant rise in flow. Therefore, typically the stroke is designed so large that 90% of the theoretical limit is achieved. For macroscopic valves, the rule of the thumb is used that the stroke should be as large as half of the radius of the valve seat. In Fig. 157, this would correspond to 50  $\mu\text{m}$  stroke which would be a suitable value, if the width of the valve seat is 20  $\mu\text{m}$ . However, even a valve seat with a width of 50  $\mu\text{m}$  (less than the diameter of a hair!) is hard to fabricate. Thus, this example shows how important it is for a microvalve to fabricate a small width of the valve seat and a large stroke.

It is also seen in Fig. 157 and from (355) that the flow is increasing with the third power of the stroke as long as the stroke is small, because the flow is limited by the gap between valve seat and valve body. Next to this stroke range, there is a more linear relationship between flow and stroke. This range is favorable for a valve which shall be employed for controlling a continuous flow. For this kind of valves, a somewhat wider valve seat is desirable which allows for a precise adjustment of the flow (if the actuator is able to provide the necessary stroke length).



**Fig. 157** Characteristic curve of a microvalve for different widths of the valve seat and theoretical limits for valve gap without feed channels and feed channels without valve gap

When the valve body is moved perpendicular to the flow the valve is called a slide valve (cf. Fig. 155b). To close the valve, the friction in the bearing of the valve body needs to be overcome by the actuator. That is, the actuator does not need to work directly against the pressure over the valve gap. The narrowest part of the flow channel is the part which is not closed by the slide and the flow is controlled by the size of that gap.

For simplicity of the calculation, it is assumed now that both the feed channel and the slide have a rectangular cross-section. The volume flow  $\Phi_S$  through the gap of a valve with cross-sectional area  $A_g$ , height  $H$ , width  $B$ , and length  $L_V$  at a pressure difference  $\Delta p_S$  is calculated corresponding to (353) and (351) as:

$$\Phi_S = v A_g = \frac{-D_h^2 H B}{32 \eta L_V} \Delta p_S. \tag{359}$$

The hydraulic diameter of the rectangular gap is calculated according to (176) on page 107:

$$D_h = \frac{4H B}{2 H + 2 B} = 2 \frac{H B}{H + B}. \tag{360}$$

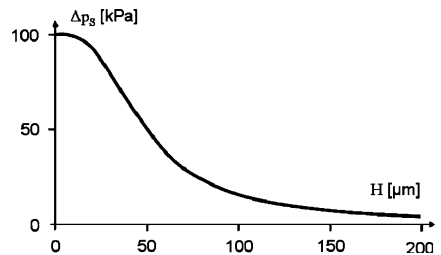
The volume flow  $\Phi_Z$  through the feed channels with height  $H_Z$ , width  $B_Z$ , overall length  $L_Z$ , and a pressure drop  $\Delta p_Z$  is calculated in the same way as for the flow through the gap of the valve. As in the case of the seat valve, the flow through the entire valve  $\Phi_F$  is the same as through the feed channels and the gap:

$$\Phi_F = \Phi_Z = \Phi_S = \frac{-H_Z^3 B_Z^3}{8 \eta L_Z (H_Z + B_Z)^2} \Delta p_Z = \frac{-H^3 B^3}{8 \eta L_V (H + B)^2} \Delta p_S. \tag{361}$$

With  $\Delta p_Z = \Delta p_V - \Delta p_S$ , it is obtained from the above equation:

$$\Delta p_S = \frac{\Delta p_V}{1 + \left( H^3 B^3 / (L_V (H + B)^2) \right) \left( L_Z (H_Z + B_Z)^2 / (H_Z^3 B_Z^3) \right)}. \tag{362}$$

Thus, the pressure difference over the gap of the slide valve is a function of the stroke  $H$ . In Fig. 158, this interrelationship is shown for a valve at which width  $B_Z$



**Fig. 158** Pressure difference over the gap of a slide valve as a function of the stroke

and height  $H_Z$  of the feed channels and the width of the slide  $B$  are all  $200 \mu\text{m}$ , the length of the gap  $L_V$  and the feed channels  $L_Z$  are  $20$  and  $500 \mu\text{m}$ , respectively, the pressure difference over the entire valve  $\Delta p_V$  is  $100 \text{ kPa}$ , and the fluid is water with a viscosity of  $1 \text{ mPa s}$ .

The flow  $\Phi_F$  through the valve as a function of the pressure difference over the entire valve is obtained now by inserting (362) into (361):

$$\Phi_F = \frac{\Delta p_V}{8 \eta \left( \left( L_V (H+B)^2 / (H^3 B^3) \right) + \left( L_Z (H_Z + B_Z)^2 / (H_Z^3 B_Z^3) \right) \right)} =: \frac{\Delta p_V}{R_{\text{fl}}}. \quad (363)$$

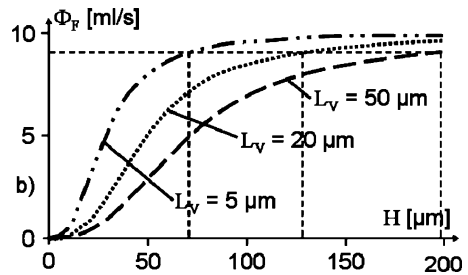
The above equation is the characteristic curve of a slide valve. The quantity  $R_{\text{fl}}$  is the flow resistance of the valve. The characteristic curve for a microvalve with the same parameters as used for Fig. 158 is shown in Fig. 159. The comparison with Fig. 157 shows that the characteristic curves are very much alike each other, but the stroke which is necessary to achieve a certain flow is for a slide valve approximately a factor 2.5 larger than for a seat valve. The larger stroke also results in a longer switching time. The necessary stroke length can be reduced by decreasing the length  $L_V$  of the gap. The limit of the length of the gap is defined by fabrication and stability issues.

As a consequence, the actuator required for a slide valve needs to achieve a larger maximum stroke  $H_{\text{max}}$ . The stroke can be reduced, if the width  $B$  of the gap is increased, while the maximum stroke length is reduced such that the hydraulic diameter remains to be constant.

According to (360), the necessary maximum stroke length for a given hydraulic diameter  $D_h$  as a function of the gap width  $B$  is:

$$H_{\text{max}} = \frac{D_h B}{2B - D_h} \Rightarrow \lim_{B \rightarrow \infty} (H_{\text{max}}) = \frac{D_h}{2}. \quad (364)$$

From the above equation, it is seen that the stroke cannot be reduced to less than half of the hydraulic diameter, even if the width of the gap is designed very large. Even if the maximum stroke length is reduced to half of the hydraulic diameter, it is still a factor of two more than for a seat valve.



**Fig. 159** Characteristic curve of a slide valve for different length of the gap



The force  $F_H$  which needs to be overcome by the actuator of a slide valve is the friction between the slide and the wall which is the slide bearing. The friction force is the product of the friction coefficient  $\mu$  and the force which presses the slide against the wall which can be estimated by the pressure difference  $\Delta p_S$  over the gap times the area  $A$  of the gap covered by the slide. The area covered by the slide is the width of the gap  $B$  times the difference of the maximum stroke length  $H_{max}$  and the actual deflection  $H$  of the slide:

$$F_H = \mu \Delta p_S A = \mu \Delta p_S B(H_{max} - H). \tag{365}$$

The pressure difference  $\Delta p_S$  over the valve gap is described by (362) and drawn in Fig. 158, but for simplicity it is approximated by a linear drop over the stroke of the slide:

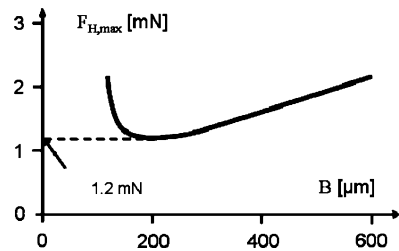
$$\Delta p_S = \left(1 - \frac{H}{H_{max}}\right) \Delta p_V \tag{366}$$

$$\begin{aligned} \Rightarrow F_H &= \mu \left(1 - \frac{H}{H_{max}}\right) \Delta p_V B(H_{max} - H) \\ &= \mu \Delta p_V B \left(H_{max} - 2H + \frac{H^2}{H_{max}}\right). \end{aligned} \tag{367}$$

The maximum actuator force  $F_{H,max}$  is required when the valve is nearly completely closed ( $H = 0$ ), because the area of the slide exposed to the pressure difference over the gap is maximum in this case. If the maximum force required to switch the valve shall be reduced by designing the optimum width  $B_{F,opt}$  of the gap and adopting the maximum stroke  $H_{max}$  such that a certain hydraulic diameter is achieved, the maximum force needs to be calculated as a function of the gap width by inserting (364) into (367) with  $H = 0$ :

$$F_{H,max} = \mu \Delta p_V B H_{max} = \mu \Delta p_V \frac{B^2 D_h}{2 B - D_h}. \tag{368}$$

The above equation is drawn in Fig. 160 as a function of the gap width for a pressure difference over the entire valve of 100 kPa, a hydraulic diameter of 200  $\mu\text{m}$ , and a friction coefficient of 0.3. The minimum of  $F_{H,max}$  is found by calculating the derivative of (368) with respect to the width and setting it equal to zero:



**Fig. 160** Maximum switching force  $F_{H,max}$  required for a slide valve as a function of the width of the valve gap  $B$

$$\frac{\partial F_{H,\max}}{\partial B} = \mu \Delta p_V \frac{2BD_h(B - D_h)}{(2B - D_h)^2} = 0 \Rightarrow B_{F,\text{opt}} = D_h. \quad (369)$$

If  $B_{F,\text{opt}}$  from the above equation is input in (364), the resulting maximum stroke  $H_{\max}$  is also the hydraulic diameter  $D_h$ . Thus, the optimum form of the slide with respect to a minimum force needed for actuation is a square. The maximum forces required for a slide valve with a square gap and for a seat valve with a square seat of the same dimensions are:

$$\text{Slide valve: } F_{H,\max} = \mu \Delta p_V B^2, \quad \text{Seat valve: } F_{H,\max} = \Delta p_V B^2. \quad (370)$$

Thus, for a slide valve, an actuator with a smaller force is sufficient than for a seat valve. The actual value of the friction coefficient is a function of the roughness and the materials of slide and wall and can vary in a comparatively large range, but it will always be much smaller than 1.

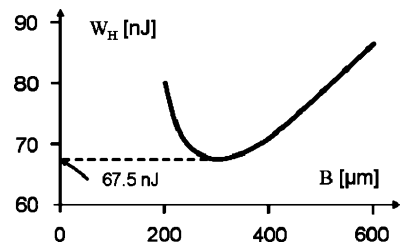
When the stroke required to switch a valve is larger for a slide valve and the force is larger for a seat valve, the question arises which valve type needs less energy for switching. The energy  $W_H$  required to switch a slide valve is found by integrating the force  $F_H$  over the stroke  $H$ . The force is available from (367):

$$\begin{aligned} W_H &= \int_0^{H_{\max}} F_H dH = \int_0^{H_{\max}} \mu \Delta p_V B \left( H_{\max} - 2H + \frac{H^2}{H_{\max}} \right) dH \\ &= \frac{1}{3} \mu \Delta p_V B H_{\max}^2. \end{aligned} \quad (371)$$

In the above equation,  $H_{\max}$  again can be expressed as a function of the width  $B$  of the valve gap at a certain hydraulic diameter  $D_h$  with (364):

$$W_H = \frac{1}{3} \mu \Delta p_V B H_{\max}^2 = \frac{1}{3} \mu \Delta p_V \frac{B^3 D_h^2}{(2B - D_h)^2}. \quad (372)$$

Figure 161 shows the energy required for switching a slide valve as a function of the width of the valve gap when the hydraulic diameter is held constant (All parameters are the same as in Fig. 160). Obviously, there is an optimum width  $B_{W,\text{opt}}$  of the gap for which the necessary switching energy is minimal. This optimum is calculated by setting zero the derivative of (372) with respect to the width and solving that equation for the width:



**Fig. 161** Switching energy  $W_H$  of a slide valve as a function of the width of the valve gap  $B$

$$\frac{\partial W_H}{\partial B} = \frac{1}{3} \mu \Delta p_V B^3 D_h^2 \frac{2B - 3D_h}{(2B - D_h)^3} = 0 \Rightarrow B_{W,opt} = \frac{3}{2} D_h. \tag{373}$$

Inserting the above result in (364) yields for the maximum stroke length:

$$H_{max} = \frac{3}{4} D_h = \frac{1}{2} B_{W,opt}. \tag{374}$$

The minimum energy required for switching the valve is now obtained from (372) by inserting  $B_{W,opt}$  for  $B$ :

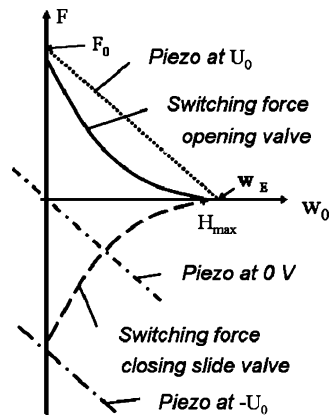
$$W_H = \frac{9}{32} \mu \Delta p_V D_h^3. \tag{375}$$

For a seat valve, the energy required for switching is calculated with (350) (page 205). If a square valve seat with an edge length  $B$  is assumed, and the stroke is a fourth of  $B$ , for the energy required for switching it is obtained:

$$W_H = \frac{1}{4} \Delta p_V D_h^3. \tag{376}$$

For the approximate calculation of the energy required for switching a seat valve, it was not taken into account that the pressure over the valve body reduces when it is lifted off the valve seat. This maybe halves the necessary energy. As a result, the comparison shows that the energy required for valve switching is nearly the same for seat valves and for slide valves.

If a normally closed seat valve as shown in Fig. 154b or in Fig. E20 (page 204 and 156, respectively) is opened, the switching force decreases during opening because the pressure drop over the valve body is reduced. The change of the switching force as a function of the stroke is schematically shown in Fig. 162



**Fig. 162** Switching force and actuator force of a piezo as a function of stroke

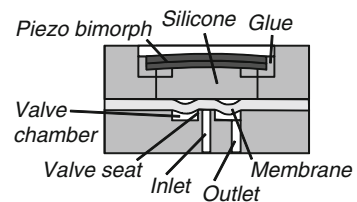
together with the characteristic curve of a piezo bimorph which could be employed as a drive (cf. Fig. 105, page 145). As seen in the figure, the forces of both the piezo and the necessary force for switching are decreasing with the deflection. Thus, the force of the piezo is sufficient to open the valve. When the valve is to be closed, the piezo is short-circuited (0 V) and its elastic force starts to push the valve body back. The more the valve body is moved towards its closed position the stronger the pressure drop over the valve body becomes and helps to close the valve, although the force of the piezo is decreasing.

If instead of a seat valve a slide valve is used, the friction force always acts against the movement of the slide. Therefore, opening of the valve is similar as for a normally closed seat valve, but the valve can be closed only if the voltage is applied to the other piezo of the bimorph.

A normally closed seat valve driven with a piezo is shown in Fig. 163 [59]. During assembly of the valve silicone was pressed into the space between piezo and valve membrane and hardened there. The pressure of the silicone deflected the piezo a little bit. Therefore, the elastic force of the piezo presses the membrane down onto the valve seat and keeps the valve closed against a pressure from the inlet. When voltage is supplied to the piezo, it bulges up for a few micrometers. The diameter of the piezo is five times larger than the diameter of the movable part of the membrane. Therefore, the membrane is lifted up 25 times more than the piezo. This way, the small deflection of the piezo is compensated and the necessary large stroke of the microvalve of at least 50  $\mu\text{m}$  is achieved.

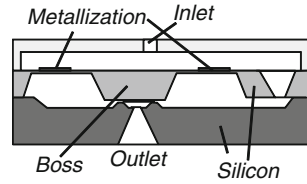
Filling the chamber between piezo and valve membrane after gluing the piezo into the valve housing, has also the advantage that fabrication tolerances are not a problem. A piezo which achieves only a stroke of a few micrometers needs to be mounted with a tolerance which is much smaller than that. By injecting the silicone into the chamber at a certain pressure, this problem is circumvented and fabrication tolerances are compensated. The direct contact of the silicone to the piezo has also the advantage that mechanical shock is damped and the risk of breaking the piezo is reduced. However, the disadvantage of this valve is the thermal expansion of the silicone which is much larger than the one of the housing. As a consequence, the voltage required to open the valve is a strong function of temperature and at low temperatures the valve may even open without voltage supplied.

One of the first microvalves developed [60] is schematically shown in Fig. 164. It employs a thermal bimaterial actuator (cf. page 161f) made of a metallization on a silicon membrane. The actuator opens the normally closed valve against the pressure from the inlet which presses the boss onto the valve seat.

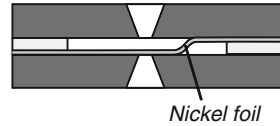


**Fig. 163** Normally closed piezo seat valve [59]

**Fig. 164** Normally closed seat valve with a bimaterial actuator [60]



**Fig. 165** Microvalve with an S-shaped electrostatic actuator [61]



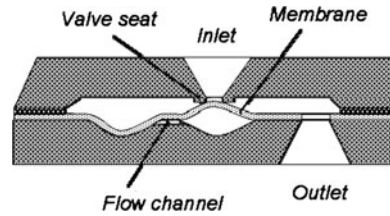
The small thermal expansion of the metallization is enhanced by the bimaterial effect. As a consequence, it may happen that the valve is already open after fabrication or closed with a comparatively large force because of residual stress in the metallization (cf. page 9f). Besides this, due to thermal stress, the valve is sensitive to temperature changes. At elevated temperatures, the valve may open if the pressure drop is too small, and at low temperatures more energy is required to open the valve.

The temperature sensitivity and the risk of residual stress affecting the valve can be decreased by a second inactive metallization as shown in Fig. E22 on page 173 [51]. However, it may be difficult to find the necessary dimensions and to ensure that fabrication meets the required tolerances.

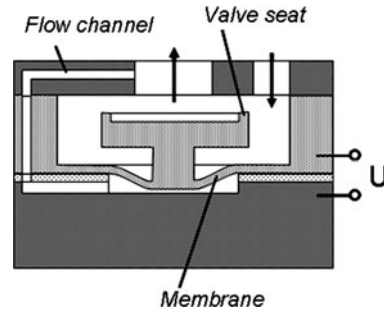
Electrostatic actuators are not sensitive to temperature changes. However, they achieve large forces only at small distances. Therefore, it is interesting to discuss the microvalve design shown in Fig. 165 [61]. A thin S-shaped foil from nickel is placed between electrically insulated electrodes at the top and the bottom of the valve chamber. When a voltage is supplied between the foil and the bottom electrode, the foil is moving to the right, because this way a larger part of the foil gets near to the bottom electrode. The bottom port is closed and the top port of the valve is opened when the foil is moving to the right. The electrostatic force is mainly working at the edge where the S-shaped foil is near to the bottom electrode or directly in touch to its insulation layer (The insulation layer is not shown in the figure). Thus, the electrostatic force can strongly attract the foil and a large stroke is achieved.

Temperature changes will have no significant effect on this type of valve because the difference in thermal expansion of nickel and the housing made of silicon only changes a little bit the slope of the S-shaped part of the foil, and no stress is generated which could affect valve switching. Another advantage is that power consumption is limited to switching. On the other hand, the long-term performance of the valve must be observed with care, because folding up and down the foil may result in wear if the foil is not thin enough, the curvature of the foil is too large, or the foil material is too brittle.

**Fig. 166** Bistable microvalve with an electrostatic actuator [62]



**Fig. 167** Microvalve which can be switched with reduced force [63]



The valve shown in Fig. 166 also employs an electrostatic actuator which is working on a short distance, while a large stroke is achieved by the design of a moving electrode [62]. In this valve, two circular membranes are placed over concave-curved electrodes. One membrane is bulging down and is in direct touch to the insulation of the electrode (The insulation is not shown in the figure), while the other membrane is bulging up and closing the inlet of the valve. The volumes between the membranes and the electrodes are connected with each other through a flow channel.

When the valve shall be opened, a voltage is applied between the membrane bulging up and the electrode below. The electrostatic force attracts the membrane mainly where it is near to the electrode until the membrane flips over to its down position. At the same time, the other membrane is pushed up by the air which is squeezed out of the volume below the downward moving membrane. The elastic forces of the two membranes keep them in their new positions and no power needs to be supplied to the valve. Switching back to the closed state of the valve is achieved by supplying voltage between the other membrane and the electrode below.

Temperature changes may affect this valve, if the membranes are not made of the same material as the housing. Besides this, outer forces acting on the housing will change the voltage required for switching. As in the case of the S-shaped actuator, long-term performance must be observed with care because wear may reduce the stability of the membranes.

The valve designs shown in Figs. 165 and 166 enlarge the stroke of an electrostatic actuator while allowing a short distance between the electrodes which yields a large force. The opposite is also possible. The force required for switching can be reduced by the design shown in Fig. 167 [63]. The pressure applied to the inlet of

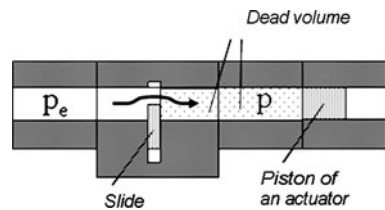
the valve is acting both on the valve body in the direction of the outlet and on the membrane pulling the valve body back from the outlet. The area on the valve body between the valve seat and the area of the membrane are adjusted such that the valve is normally closed. If it shall be opened, a voltage is applied between the membrane and an electrode on the lower part of the housing. Thus, the membrane pulls the valve body away from the outlet. The distance between membrane and counter electrode on the housing needs to be comparatively large in this case, because a significant stroke is required to open the valve. However, the necessary force is reduced because the pressure from the inlet is also acting on the membrane, and, therefore, is balanced a bit.

If the membrane of the valve shown in Fig. 167 is not made of the same material as the housing, temperature changes will have side effects on the valve because the residual stress of the membrane will change and the force required to deflect the membrane is a function of temperature. Similar effects will occur when outer forces are acting on the housing. Another important issue is the tolerance during fabrication which needs to be small enough to ensure that the valve is normally closed and can be opened by the weak force of the electrostatic actuator. Besides this, when only a small force is required to open the valve, it could be switched also by an impact on the housing.

The time between applying the switching signal to a valve and the rising of the pressure or the flow behind the valve is an important parameter for many applications such as pneumatic and hydraulic controls. Before a system behind the valve can react, the pressure needs to be built up between the valve seat and the system. In a hydraulic system, the pressure propagates with the speed of sound, and, since microvalves are very small, this is not an issue. Switching time in a hydraulic system is just a function of the time required to move the valve body.

In a pneumatic system, this is different, because a certain flow of gas into the volume between the valve seat and the following system is required until the full pressure is arriving (cf. Fig. 168). The volume between valve seat and the point where the pressure rise is required is called the *dead volume*. The interrelationship between volume and pressure of a gas in a certain volume in most cases is described sufficiently well by the ideal gas law (295 on page 167). From this law, it is found that the gas volume  $dV$  which needs to enter the dead volume  $V_T$  to raise the pressure  $p$  by  $dp$  is:

$$dV = \frac{V_T}{p} dp. \tag{377}$$



**Fig. 168** Dead volume between a valve and an actuator

The volume flow  $\Phi_F$  through a valve is calculated from the pressure difference  $\Delta p_V$  and the flow resistance  $R_{fl}$  of the valve (cf. (355) and (363) on pages 207 and 210, respectively). Together with the above equation it is obtained:

$$\Phi_F = \frac{\Delta p_V}{R_{fl}} = \frac{dV}{dt} = \frac{V_T}{p} \frac{dp}{dt} \Rightarrow \frac{dp}{dt} = \frac{p \Delta p_V}{V_T R_{fl}} = \frac{p(p_e - p)}{V_T R_{fl}}. \quad (378)$$

Separation of variables and integration allows calculating the pressure in the dead volume as a function of the time after a sudden opening of the valve:

$$dt = \frac{V_T R_{fl}}{p(p_e - p)} dp \Rightarrow \int_0^t dt' = \int_p^{p_e} \frac{V_T R_{fl}}{p' (p_e - p')} dp' \quad (379)$$

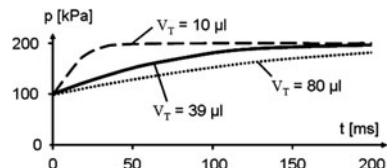
$$\begin{aligned} &= t = -\frac{V_T R_{fl}}{p_e} \ln \left( \frac{(p_e - p)(p_e - \Delta p_V)}{p \Delta p_V} \right) \\ &\Rightarrow p = \frac{(p_e - \Delta p_V)p_e}{\Delta p_V e^{-(p_e/(V_T R_{fl}))t} + p_e - \Delta p_V}. \end{aligned} \quad (380)$$

The pressure rise calculated with (380) for a microvalve with a fluidic resistance of 223 Pa s/ $\mu$ L, a pressure drop of 100 kPa, and a pressure supply of 200 kPa at the entrance of the valve is shown in Fig. 169 as a function of the dead volume. This clearly demonstrates the effect of the dead volume on the rise time of the pressure. The smaller the dead volume is, the quicker the pressure behind the valve will change when the valve is opened.

According to (380), the pressure difference over the valve reduces by a factor  $1/e$  ( $\approx 1/2.7$ ) in the pressure rise time  $\tau_p$ :

$$\tau_p = \frac{V_t R_{fl}}{p_e}. \quad (381)$$

As a consequence of the above equation, the product of the dead volume  $V_t$  and the fluidic resistance  $R_{fl}$  needs to be minimized when a quick response time of a pneumatic system is required. That is, fluid channels should be designed as short and as wide as possible. This means also that the flow path between the valve seat and the outlet of a microvalve should be designed as short and as wide as possible.



**Fig. 169** Pressure rise in dead volume calculated with (380)



The fluidic resistance of a capillary is found from (178) on page 108. In Table 11 on page 110, there are listed some fluidic resistances of capillaries which are often employed in microtechnique. For a channel with a circular cross-section, the pressure rise time  $\tau_p$  according to (381) is:

$$\tau_p = \frac{8 \eta L_K^2}{p_e R_{Ka}^2}. \quad (382)$$

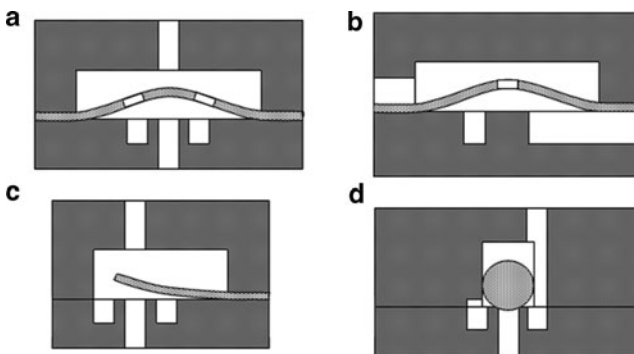
The above equation shows that scaling up or scaling down all linear dimensions of the channel does not change the pressure rise time. Thus, the switching time of a pneumatic microvalve could be expected to be the same as for a macroscopic valve. However, microvalves can achieve shorter switching times because the stroke of the valve body is smaller, and, therefore, this smaller distance can be covered in a shorter time.

If only a back flow is to be avoided, a valve does not need an actuator and a *passive valve* can be employed. The main application of passive microvalves are micropumps (see next chapter).

The demands on passive microvalves are similar as on active ones. A small flow resistance is required in forward direction, the valve needs to be tight in backward direction and it should open and close as quickly as possible. Of course, as always, the fabrication cost shall be low and the performance of the valve should not be a function of parameters such as temperature and humidity.

Typical designs of microvalves are shown in Fig. 170. The valve body is a beam clamped at one or both ends, a membrane, or a loose part such as a sphere or a membrane. Besides the expected characteristic of the valve, fabrication issues need also to be taken into account for the design. Maybe, for an application, a certain design promises the largest flow but with the available fabrication equipment another design is favorable, and therefore the best choice.

The flow through an open passive microvalve can be calculated with (355) (on page 207), but the stroke needs to be calculated from the pressure difference over



**Fig. 170** Passive microvalve designs employing as valve body a membrane (a, b), a beam clamped at one end (c), and, a loose sphere (d)

the valve seat. If the stroke is not too large, the pressure difference over the valve seat will be nearly as large as the one which moves the valve body off.

If an elastic beam or membrane is employed as valve body, the deflection  $w_0$  (now representing the stroke  $H$ ) can be calculated with an equation from Table 6 on page 80. For example, for a beam with Young's modulus  $E_B$ , length  $L_B$ , width  $b_B$ , and thickness  $d_B$  clamped at one end and a circular feed channel with radius  $R_Z$ , it is found:

$$H = 4 \frac{L_B^3}{E_B b_B d_B^3} F = 4 \frac{L_B^3 \pi R_Z^3}{E_B b_B d_B^3} \Delta p_S. \quad (383)$$

The above equation shows that the stroke of a passive valve employing as valve body a beam clamped at one end is approximately a linear function of the pressure over the valve seat.

If the valve seat is assumed annular and its radius is the same as the one of the feed channel, the pressure drop over the valve seat  $\Delta p_S$  can be expressed by the pressure difference over the entire valve  $\Delta p_V$  with (357) (page 208). As a result, the interrelationship between the stroke of the beam and the pressure difference over the valve becomes:

$$\Delta p_V = \frac{E_B b_B d_B^3}{4\pi L_B^3 R_Z^3} \left( 2 \frac{L_Z}{L_V} H^3 + R_Z^3 \right) H. \quad (384)$$

If the stroke (and the pressure drop over the valve) is small, the first term in the parenthesis of the above equation can be neglected compared with the second one, and the stroke can be expressed as a function of the pressure difference over the valve. This shows that the stroke is a linear function of a small pressure difference:

$$H = \frac{4\pi L_B^3 R_Z^3}{E_B b_B d_B^3} \Delta p_V. \quad (385)$$

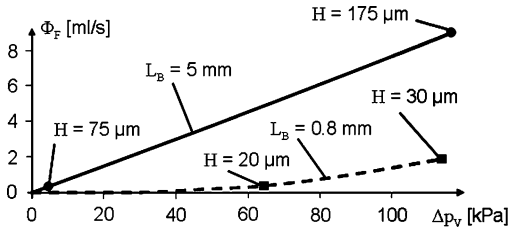
If the pressure difference (and the stroke of the beam) is large, the second term in the parenthesis of (384) can be neglected resulting in:

$$H = \sqrt[4]{\frac{2\pi L_B^3 R_Z^5 L_V}{E_B b_B d_B^3 L_Z}} \Delta p_V. \quad (386)$$

The above equation shows that the stroke is only a weak function of the pressure difference over the valve when the stroke is large.

The characteristic curve of a passive valve is the flow through the valve as a function of the pressure difference. This curve cannot be calculated in general, because the stroke according to (384) cannot be expressed as a simple function of the pressure difference. But it is possible to calculate for a given stroke  $H$  the corresponding pressure difference and with this pressure drop and the stroke the

**Fig. 171** Flow through passive valves equipped with beams, 5 and 0.8 mm in length, respectively, as valve bodies calculated with (384) and (355)



volume flow according to (355) (page 207). This way the characteristic curve of a passive microvalve with feed channels with a length of 500  $\mu\text{m}$ , a radius of 100  $\mu\text{m}$ , and both radius and width of the valve seat of 100  $\mu\text{m}$  was calculated and shown in Fig. 171. For this calculation, a beam was assumed as valve body with a width of 200  $\mu\text{m}$ , a thickness of 20  $\mu\text{m}$ , and a Young’s modulus of 120 GPa, and the viscosity of the flow medium was 1 mPa s (water).

Figure 171 shows that for a large stroke  $H$  (which is obtained with a long beam) the flow  $\Phi_F$  through the passive valve is a linear function of the pressure difference  $\Delta p_V$  over the valve. That is what has to be expected from (355): When the stroke is large, the first term in the parenthesis in the denominator can be neglected, the flow is only a function of the geometry of the feed channels and is proportional to the pressure difference. This is the maximum flow which can be achieved for a passive valve. The maximum flow can be enhanced by widening and shortening the flow channels and choosing their cross-section as close as possible to a circular one (resulting in a larger hydraulic diameter).

If the stroke is small, the second term in (355) can be neglected compared with the first one and the flow is rising with the third power of the stroke. According to (385), the small stroke of a beam is proportional to the pressure difference over the valve. Inserting (385) into (355) yields:

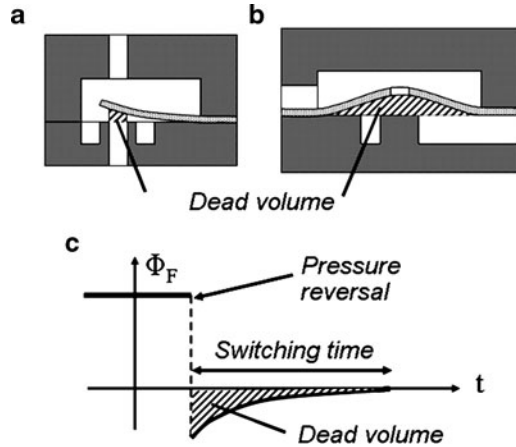
$$\Rightarrow \Phi_F = 8 \frac{\pi^3 L_B^9 R_Z^6 U_s}{\eta E_B^3 b_B^3 d_B^9 L_V} \Delta p_V^4. \tag{387}$$

The above equation shows that for a small stroke the flow through a passive valve with a beam clamped at one end as a valve body is a function of the fourth power of the pressure drop over the valve.

Every passive valve will show a small stroke when the pressure difference over the valve is small enough, and nearly every passive valve can show a large stroke when the pressure difference is large enough (if it is designed properly and does not break). Thus, the designer will try to choose the materials and dimensions such that a small pressure difference results in a large stroke and consequently in a large flow through the valve.

The larger a valve is, the more flow is achieved; but there is a reason not to increase the dimensions of a passive valve too much: the dead volume. If the direction of the pressure difference over the valve is reversed, the flow through

**Fig. 172** Dead volume of valves (a) and (b) and reverse flow due to dead volume (c)



an ideal passive valve should stop immediately. In the reality, however, the fluid volume between valve body and valve seat in the open position of the valve can flow in backward direction resulting in a reverse flow as shown in Fig. 172.

The reverse flow of the dead volume takes some time because friction in the feed channels hinders the flow. Therefore, the valve is not completely closed before the back flow stopped and the duration of the back flow is called the switching time of the valve. Obviously a short switching time is desirable.

It is difficult to calculate the dead volume of a valve exactly, because the flow around the valve body would need to be taken into account. However, as an approximation, the dead volume can be calculated as the volume between the valve body of the open valve and the valve seat as shown in Fig. 172a, b. For a beam or a loose valve body and a circular valve seat, the dead volume  $V_V$  can be approximated by the cylinder with the cross-section of the feed channel with radius  $R_Z$  and the distance  $H$  between valve body and valve seat:

$$V_V = \pi R_Z^2 H. \quad (388)$$

If the valve body is a membrane as shown in Fig. 172b, the dead volume can be approximated by the volume of the deflected membrane. For a circular membrane with radius  $R_M$  according to (303) on page 170, this results in:

$$V_V = \frac{1}{2} \pi R_M^2 H. \quad (389)$$

Since the radius of the membrane is always larger than the radius of the feed channel, a valve with a membrane has a comparatively large dead volume. This does not necessarily mean that passive membrane valves are the worse solution, because, e.g., a membrane may easier be fabricated than a beam which more easy can be broken.

When a passive valve is closing, its dead volume needs to be displaced either around the valve body, through the orifice of a valve membrane, or back into the feed channel. Which of these cases occurs is a function of the elastic force pulling the valve body down onto the valve seat, the friction of the flow around or through the valve body, and the friction in the feed channel. In each of these cases, some time is needed to close the passive valve. The calculation of that time may be difficult.

The designer of a passive valve will try to reduce the elastic force pulling the valve body back onto the valve seat because this force would need to be overcome to open the valve and it is desirable to open the valve even with a small pressure difference. Therefore, in many cases, it may be assumed that the elastic force can be neglected and the dead volume is pushed in backward direction only with the pressure difference over the valve. Thus, the backflow  $\Phi_b$  can be estimated with (178) (on page 108) and the flow resistance  $R_{f,z}$  of the feed channels of the valve which may be found in Table 11 (page 110). The dead volume  $V_V$  is approximately the product of the backflow and the switching time  $t_V$  of the valve:

$$V_V = \Phi_b t_V = \frac{\Delta p_V}{R_{f,z}} t_V \Rightarrow t_V = \frac{R_{f,z} V_V}{\Delta p_V}. \quad (390)$$

## Exercises

### *Problem 30*

You shall design a microvalve which is able to switch pressure differences of up to 1 MPa. The valve shall be switched with a membrane which is equipped with an electrostatic actuator. The radius of the membrane shall be 5 mm, the minimal distance between membrane and counter electrode which can be fabricated accurately is 5  $\mu\text{m}$  and a maximum voltage of 200 V can be applied. The length of the flow channel over the valve seat is 50  $\mu\text{m}$ .

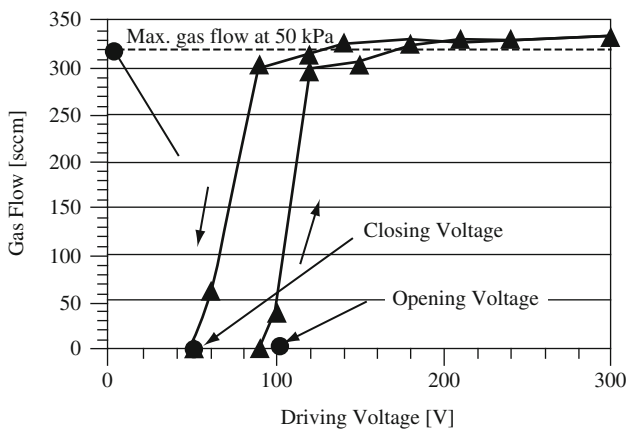
- What is the maximum pressure which can be generated with the electrostatic actuator when the membrane is not deflected?
- As a consequence of the result of (a), you decide to build a pressure balanced valve according to Huff et al. (cf. Fig. 167). How large may the fabrication tolerance of the radii of the two membranes in the valve be to make sure that the valve can still work against the desired pressure?
- Assume that the valve is employed in a pneumatic system with feed channels with a radius of 100  $\mu\text{m}$  and an overall length of 10 cm. How many percent of the pressure drop over the entire system is lost at the valve?
- What is the volume flow through the valve for water (viscosity = 1 mPa s), with the maximum pressure difference of 1 MPa?

- (e) How large is the volume flow through the completely opened valve if a fabrication tolerance of  $2\ \mu\text{m}$  is achieved, and the stroke is enhanced such that the actuator force is still sufficient?
- (f) What could be the reason why the valve should not be designed as described in (e)?

### Problem 31

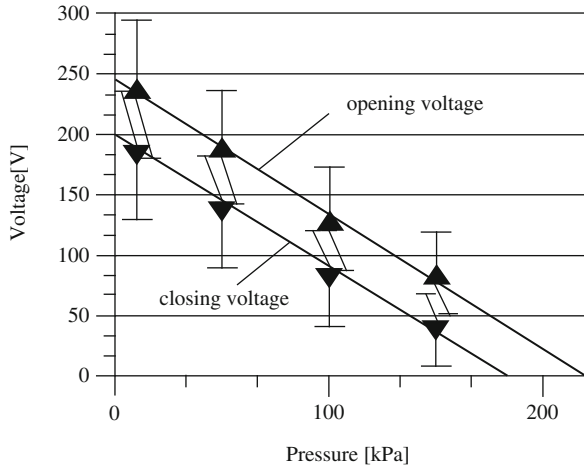
The microvalve shown in Fig. 163 is driven with a circular disk of PZT ceramic which is glued onto a steel disk. Figure E25 shows the measured nitrogen flow through the valve as a function of the voltage applied at the actuator. (The unit sccm means “Standard  $\text{cm}^3$  per minute” according to  $\text{cm}^3/\text{min} = 16.7\ \mu\text{L/s}$  at a pressure of 101.3 kPa and a temperature of  $22^\circ\text{C}$ .)

- (a) How long was the length of the feed channels to the valve, if they showed the same diameter ( $200\ \mu\text{m}$ ) as the orifices and feed channels in the valve?  
Hint: Compare the graph with Fig. 157 on page 208 and find a way to modify (355).
- (b) In Fig. E26, there are shown the voltages at which the valve was opening and closing, respectively, as a function of the over pressure applied to the inlet.
- (c) What was the residual pressure of the silicone?
- (d) With what force pulls the actuator at the silicone when the valve is opened without pressure at the inlet?
- (e) How many force needs to be applied to overcome material straining when the valve is opened? Assume that the maximum force of the actuator is available for that.



**Fig. E25** Nitrogen flow through the valve shown in Fig. 163 as a function of the voltage supplied to the piezo. (Reprinted from [59] with permission from Elsevier)

**Fig. E26** Voltage at which the valve shown in Fig. 163 is opening (*triangles*) and closing (*inverted triangles*), as a function of the pressure applied at the inlet. (Reprinted from [59] with permission from Elsevier)



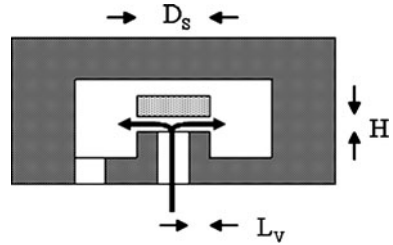
Thickness of PZT disk	200 $\mu\text{m}$
Young's modulus of PZT disk	67 GPa
Piezoelectric modulus $d_{31}$ of PZT disk	190 $\mu\text{m/V}$
Thickness of steel disk	100 $\mu\text{m}$
Young's modulus of steel disk	190 GPa
Diameter of the actuator	10 mm
Viscosity of nitrogen	16.2 $\mu\text{Pa s}$
Diameter of inlet, outlet, and feed channels in the valve	200 $\mu\text{m}$
Length of the feed channels in the valve	3 mm
Transmission ratio of the silicone hydraulic	1:25
Width of the sealing of the valve seat	200 $\mu\text{m}$
Diameter of the movable part of the valve membrane	2 mm

### Problem 32

In Fig. E27, there is shown the schematic drawing of an active microvalve with an annular valve seat. The diameter of the seat  $D_S$  is 0.4 mm and the length  $L_V$  of the valve gap is 20  $\mu\text{m}$ . The feed channels are 5 cm long and show a radius of 0.2 mm. The stroke length  $H$  is 0.1 mm. Water with a viscosity of 1 mPa s is flowing through the valve and the pressure difference over the valve is 500 kPa.

- Please calculate the maximum volume flow through the valve.
- The actuator can open the valve with a velocity of 2 mm/s. What is the switching time of this valve?
- How can the switching time be reduced to 50% if the same actuator is used? What is the consequence of that change on the performance of the valve?
- Since you need a valve with a reduced switching time but the same performance as in (a), you decide to make a redesign. The cross-section of the feed channels and the actuator are not allowed to be changed. What possibilities are

**Fig. E27** Schematic drawing of an active microvalve



available for you? Suggest a design for the case that only annular valve seats can be fabricated. What is the disadvantage of this solution?

### **Problem 33**

The bistable microvalve shown in Fig. 166 (page 216) was designed for an implantable drug dosing system. The diameters of the two membranes are 2 mm and their maximum deflection is 5 mm. A force of 1 mN is needed to snap a membrane to the other side.

The pressure in the drug reservoir and the tissue in which the drug shall be delivered are 20 and 10 kPa, respectively. The inlet of the valve has a square cross-section with an area of  $0.01 \text{ mm}^2$  and a length of  $10 \text{ }\mu\text{m}$ . The viscosity of the drug is  $1.05 \text{ mPa}\cdot\text{s}$ .

Neglect, in the following, the compressibility of the air between the membranes and the electrodes and, as an approximation, assume that all forces act on the centers of the membranes.

- Calculate the voltage which is required to open the valve, if there is no pressure difference over the valve. Assume as an approximation that the membrane is attracted by an electrostatic force which is as large as the one attracting an annular area at the circumference of the membrane which is  $200 \text{ }\mu\text{m}$  wide and has a distance of  $0.5 \text{ }\mu\text{m}$  from the electrode.
- Up to which pressure in the drug reservoir does the membrane remain to be closed? Assume that a gap between membrane and valve seat is developing already at a force which is 25% of the elastic force of the membrane.
- How long needs the valve to be open to deliver  $9.1 \text{ }\mu\text{L}$ ? Assume that the membrane of the open valve does not reduce the drug flow and neglect the effect of feed channels in the system compared with the effect of the inlet orifice.

### **Problem 34**

On page 218f, the time is calculated which is needed for the pressure rise in a capillary with circular cross-section. What is the corresponding equation for a rectangular capillary?



### Problem 35

The time between the switching pulse and the pressure rise behind the microvalve shown in Fig. 163 (page 214) is shown in Fig. E28. For this measurement at the end of the valve, there was connected a 50-mm long hose with an inner diameter of 900  $\mu\text{m}$ . At the end of the hose, a pressure sensor with a dead volume of 15  $\mu\text{L}$  was mounted. The pressure difference over the valve was 50 kPa over environmental pressure and nitrogen was used as the fluid.

- Please calculate the dead volume of sensor and hose together.
- The dead volume of the valve is 0.33  $\mu\text{L}$ . Please calculate the fluidic resistance of the hose and the valve.
- How quick would the pressure rise be in the hose and the sensor alone, if the microvalve had neither a flow resistance nor a dead volume?
- How quick could the pressure rise be, if it would be measured directly at the outlet of the valve and the actuator would react infinitely quick? Assume that the pressure rise is a factor  $e$  (approximately 2.7) longer than the pressure rise time.
- How quick could the pressure rise time of the entire system (valve with hose and sensor) be if the actuator would react infinitely quickly?
- How could the switching time be further enhanced?
- What could have been done better with the measurement described above?

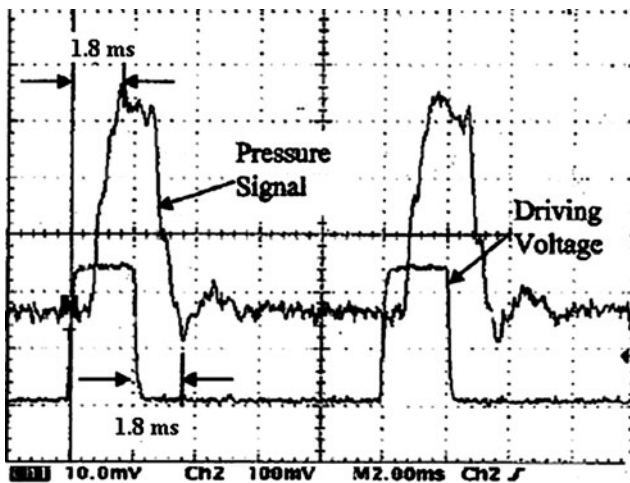


Fig. E28 Switching pulse and pressure rise measured behind the microvalve shown in Fig. 163. (Reprinted from [59] with permission from Elsevier)

### ***Problem 36***

A slide valve shall be mounted into a microchannel which is 50  $\mu\text{m}$  high and 100  $\mu\text{m}$  wide. With this microchannel a liquid similar as water shall be dosed.

- (a) How large needs the stroke of the slide be to achieve 90% of the maximum flow, if the slide is moving in the direction of the channel height?

Hint: It helps very much if a spreadsheet program is used to solve this problem.

- (b) What is the minimum force the actuator needs to provide?
- (c) Assume that the actuator consists of two circular piezo disks with the same thickness such as it is shown in Fig. 110 on page 149. What diameter need these piezo disks have to produce a stroke of 33  $\mu\text{m}$ ?

### ***Problem 37***

A pressure difference of 10 kPa is applied over a passive microvalve equipped with a beam fixed at both ends.

- (a) How large is the stroke of that beam, if there is no residual stress when the valve is closed?

Hint: It is best to use a programmable pocket calculator or a spreadsheet program to solve this problem.

- (b) What is the flow through the valve?
- (c) How could the design of the valve be improved?

Viscosity of nitrogen	16.5 $\mu\text{Pa s}$
Diameter of inlet, outlet, and feed channels of the valve	200 $\mu\text{m}$
Length of feed channels in the valve	1 mm
Width of the sealing area of the valve seat	200 $\mu\text{m}$
Length of the bar	8 mm
Width of the bar	600 $\mu\text{m}$
Thickness of the bar	20 $\mu\text{m}$
Young's modulus of the bar	120 GPa
Residual stress of the bar	0 MPa

# Micropumps

The main applications for which micropumps have been developed are drug delivery and taking and transporting of samples for analysis. It is much easier to build a micropump than a microvalve, and much more publications exist on micropumps than on microvalves. The problem with micropumps is that up to now there is no large market for them. Many tasks which could be done by micropumps can be done by even easier devices. For example, drug delivery can be done by an infusion bottle hanged up.

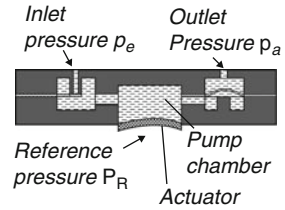
Typical specifications for micropumps are the maximum flow rate and counter pressure, accuracy of dosing and flow rate, delivery of gases and liquids, self priming, tolerance against particles and bubbles in the fluid, life time, small dimensions and weight (with power supply and electronics), small energy consumption, small temperature change and pulsation of the fluid, chemical and biological inertness, and low price.

Most micropumps are reciprocating pumps, i.e., an actuator is continuously reducing and enlarging the volume of a pump chamber. Passive valves at the inlet and the outlet cause the in and out flow to go only into one direction. Figure 173 shows the principle of such a pump. The pressure at the inlet  $p_e$  is smaller than the one at the outlet  $p_a$  and both may be different than the reference pressure  $p_R$  (typically from the environment), which is acting on the backside of the actuator.

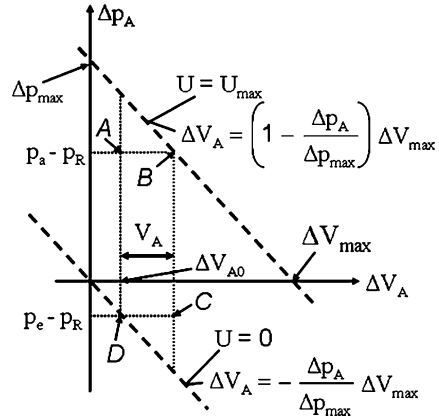
The performance of a micropump certainly is a function of its actuator. There are only a few actuator principles such as the lateral capacitive force (cf. page 132), the evaporation of a liquid in a closed volume (cf. page 171), and an external pneumatic or hydraulic drive for which the force is not a function of the stroke. Nearly all other actuators show a characteristic curve which is similar to the one of a piezo or bimaterial actuator. For example, compare the characteristic curves in Fig. 125a (page 169), Fig. 123b (page 166), and Fig. 101 (page 140) or Fig. 118 (page 162). Therefore, the following discussion assumes a piezo actuator.

For the application of an actuator for a micropump, it is best to draw the characteristic curve as the generated pressure difference  $\Delta p_A$  over the volume displacement  $\Delta V_A$  as shown in Fig. 174. The characteristic curve is defined by the maximum pressure difference  $\Delta p_{\max}$  which is achieved if the actuator is hindered from deflecting and the maximum volume displacement  $\Delta V_{\max}$  which is

**Fig. 173** Schematic cross-section of a reciprocating micropump



**Fig. 174** Characteristic curve of a typical micropump actuator



obtained when there are no outer forces acting. The characteristic curve is described by the following equation:

$$\Delta p_A = \Delta p_{\max} \left( 1 - \frac{\Delta V_A}{\Delta V_{\max}} \right) \Rightarrow \Delta V_A = \Delta V_{\max} \left( 1 - \frac{\Delta p_A}{\Delta p_{\max}} \right). \quad (391)$$

The above equation is right when the maximum voltage  $U_{\max}$  is applied to the piezo. If the piezo is short-circuited, the curve is shifted such that it touches the origin, and the characteristic curve is described by:

$$\Delta p_A = - \frac{\Delta p_{\max}}{\Delta V_{\max}} \Delta V_A \quad \text{or} \quad \Delta V_A = - \frac{\Delta V_{\max}}{\Delta p_{\max}} \Delta p_A. \quad (392)$$

When the actuator is powered, it raises the pressure in the pump chamber until it becomes larger than the pressure at the outlet ( $A$  in Fig. 174). At this point of the pump cycle, the actuator overcomes the pressure difference between the outlet and the reference pressure ( $p_a - p_R$ ). The outlet valve opens and the actuator pushes fluid out through that valve. While doing so, the pressure which can be generated by the actuator is reducing with the volume which has been displaced, until the actuator is no longer able to overcome the pressure difference ( $B$ ). Then the actuator is moving back (discharged if it is a piezo) and the lower characteristic curve is valid. The pressure in the pump chamber by the actuator now can be made smaller than the inlet pressure. Therefore, the outlet valve is closing, the inlet valve opens and

the pressure drops a bit below the pressure at the inlet. At this point of the pump cycle (C), the pressure acting on the actuator is the difference between inlet pressure and reference pressure ( $p_e - p_R$ ). Fluid is taken in through the inlet valve until the actuator can no longer provide enough underpressure (D). Then, a new pump cycle is started by powering the actuator again.

The volume  $V_A$  which is pumped during one cycle is found in Fig. 174 as the distance between the vertical dotted lines. It is calculated from the difference of the displaced volume  $\Delta V_A$  at B [(391) with  $\Delta p_A = p_a - p_R$ ] and at D [(392) with  $\Delta p_A = p_e - p_R$ ]:

$$\begin{aligned} V_A &= \Delta V_{\max} \left( 1 - \frac{p_a - p_R}{\Delta p_{\max}} \right) - \Delta V_{\max} \left( - \frac{p_e - p_R}{\Delta p_{\max}} \right) \\ &= \Delta V_{\max} \left( 1 - \frac{p_a - p_e}{\Delta p_{\max}} \right). \end{aligned} \quad (393)$$

The displaced volume  $\Delta V_A$  at D in Fig. 174 is denoted here as  $\Delta V_{A0}$  for later use in this chapter:

$$\Delta V_{A0} := \Delta V_{\max} \left( - \frac{p_e - p_R}{\Delta p_{\max}} \right). \quad (394)$$

The pump cycle between the characteristic curves of the actuator and the volume displaced per pump cycle do not contain the information how long a pump cycle lasts. This is a function of the flow through the valves. The flow  $\Phi_a$  through the outlet valve with flow resistance  $R_{fl}$  is calculated with (355) (page 207). The pressure difference  $\Delta p_V$  over the valve is with (391):

$$\Delta p_V = p_R + \Delta p_A - p_a = p_R - p_a + \Delta p_{\max} \left( 1 - \frac{\Delta V_A}{\Delta V_{\max}} \right) \quad (395)$$

$$\Rightarrow \Phi_a = \frac{-\Delta p_V}{R_{fl}} = \frac{p_R - p_a + \Delta p_{\max} \left( 1 - \frac{\Delta V_A}{\Delta V_{\max}} \right)}{R_{fl}} = \frac{\partial \Delta V_A}{\partial t}. \quad (396)$$

The flow through the outlet valve as described by the above equation is equal to the derivative of the volume displacement of the actuator. As a consequence, the volume displacement as a function of time can be calculated by separation of variables and integration of the above equation:

$$\Rightarrow \int_0^t \frac{dt'}{R_{fl}} = \int_{\Delta V_{A0}}^{\Delta V_A} \frac{d\Delta V_A'}{p_R - p_a + \Delta p_{\max} - \frac{\Delta p_{\max}}{\Delta V_{\max}} \Delta V_A'} \quad (397)$$

$$\Rightarrow \frac{t}{R_{fl}} = - \frac{\Delta V_{\max}}{\Delta p_{\max}} \ln \left( \frac{p_R - p_a + \Delta p_{\max} - \frac{\Delta p_{\max}}{\Delta V_{\max}} \Delta V_A}{p_R - p_a + \Delta p_{\max} - \frac{\Delta p_{\max}}{\Delta V_{\max}} \Delta V_{A0}} \right) \quad (398)$$

$$\Rightarrow \Delta V_A = \left( \frac{p_R - p_a}{\Delta p_{\max}} + 1 \right) \Delta V_{\max} - \left( \left( \frac{p_R - p_a}{\Delta p_{\max}} + 1 \right) \Delta V_{\max} - \Delta V_{A0} \right) e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t}{R_{fl}}} \quad (399)$$

In the above equations, it was assumed that the flow resistance  $R_{fl}$  of the valve is not a function of time. In general this is not true, but the valves of a pump should be designed such that the stroke for most of time is large to allow a flow which is limited only by the cross-section of the feed channels (cf. Fig. 157 on page 208). Thus, a constant flow resistance of the valves appears to be a suitable approximation.

Inserting (394) into (399) yields:

$$\Delta V_A = \Delta V_{\max} \left[ \left( \frac{p_R - p_a}{\Delta p_{\max}} + 1 \right) - \left( \frac{p_e - p_a}{\Delta p_{\max}} + 1 \right) e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t}{R_{fl}}} \right] \quad (400)$$

The flow through the outlet valve is now obtained by either inserting the above equation into (396) or by differentiating the above equation with respect to time:

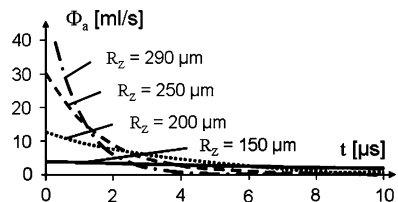
$$\Phi_a = \frac{\Delta p_{\max} - \Delta p_P}{R_{fl}} e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t}{R_{fl}}} \quad (401)$$

In the above equation, the pressure drop over the pump was denoted as  $\Delta p_P = p_a - p_e$ .

The flow through the outlet valve as a function of time and the radius  $R_Z$  of the feed channels and the valve seat calculated with (401) is shown in Fig. 175. A pump actuator was assumed which is able to generate a maximum pressure difference  $\Delta p_{\max}$  of 20 kPa and a maximum volume displacement  $\Delta V_{\max}$  of 100 nL. The pressure difference over the pump  $\Delta p_P$  was 10 kPa and water (viscosity = 1 mPa s) was used as the fluid. The flow resistance  $R_{fl}$  of the valve was calculated with (355) (page 207) for a passive valve with circular cross-section and feed channels and a large stroke (the first term in the large parenthesis in the denominator is neglected). The length of the feed channels was assumed to be 500  $\mu\text{m}$ .

In Fig. 175, the flow through the outlet valve of the pump is increasing with the size of the valve and the feed channels. This is due to the decreasing friction of the fluid in the feed channels. Thus, it appears as if it would be best to design as large feed channels as possible. This is not true, because the dead volume of the valves has a major effect and has not been taken into account in the above calculation.

**Fig. 175** Flow through the outlet valve of a micropump as a function of time and the radius of the feed channel. The dead volume of the valve is not taken into account



The volume delivered with each pump stroke is not the volume  $V_A$  shown in Fig. 174, but the actuator needs to displace the dead volume of the passive valves before fluid is leaving the outlet valve.

A certain pressure difference is necessary to open the outlet valve. Therefore, the point  $A$  in Fig. 176 is a bit above the pressure  $(p_a - p_R)$  which is necessary to overcome the pressure at the outlet. This is not taken into account here. Instead Fig. 176 shows at  $A$  the effect of stiction of a valve what may occur.

Then fluid is displaced out of the pump chamber and leaves through the outlet valve. When point  $B$  is reached, no more fluid can be displaced by the actuator and it is necessary to discharge the piezo and switch to the lower characteristic curve. As a consequence, the outlet valve is closing and its dead volume  $V_a$  is flowing backward.

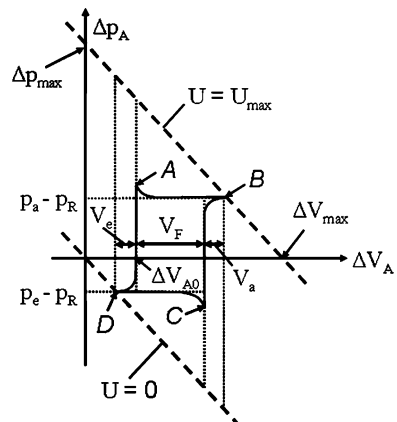
The pressure in the pump chamber is decreasing now until it is below the pressure at the inlet  $(p_e - p_R)$  and is able to open the inlet valve ( $C$  in Fig. 176). The actuator is moving back then until it is no longer able to produce a pressure which is less than the pressure at the inlet ( $D$ ). The piezo is charged again and the inlet valve is closing. As a result, the dead volume  $V_e$  of the inlet valve is displaced by the actuator in backward direction. After that the next pumping cycle is starting.

The volume  $V_F$  which is delivered by one pump cycle is the distance  $V_A$  between the dotted lines in Fig. 174 (page 230) [(calculated with (393)] minus the dead volumes of inlet  $V_e$  and outlet  $V_a$  valve:

$$V_F = V_A - V_e - V_a = \Delta V_{\max} \left( 1 - \frac{p_a - p_e}{\Delta p_{\max}} \right) - V_e - V_a \quad (402)$$

The volume displaced by the actuator as a function of time can now be calculated with (399), if the dead volume of the inlet valve  $V_e$  is added to  $\Delta V_{A0}$  in (394) (page 231):

$$\Delta V_A = \Delta V_{\max} \left[ \left( \frac{p_R - p_a}{\Delta p_{\max}} + 1 \right) - \left( \frac{p_e - p_a}{\Delta p_{\max}} + 1 - \frac{V_e}{\Delta V_{\max}} \right) e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t}{R_{fl}}} \right] \quad (403)$$



**Fig. 176** Characteristic curve of a typical micropump actuator

The flow  $\Phi_a$  through the outlet valve as a function of time is obtained when the above equation is inserted into (396) or by differentiating the above equation with respect to time:

$$\Phi_a = \frac{(1 - (V_e/\Delta V_{max}))\Delta p_{max} - \Delta p_p}{R_{fl}} e^{-\frac{\Delta p_{max}}{\Delta V_{max}} \frac{t}{R_{fl}}} \quad (404)$$

The dead volume of the inlet valve was calculated according to (388) (page 222), and the stroke was assumed to be as large as half of the radius of the feed channel:

$$V_e = \pi R_Z^2 H = \frac{1}{2} \pi R_Z^3. \quad (405)$$

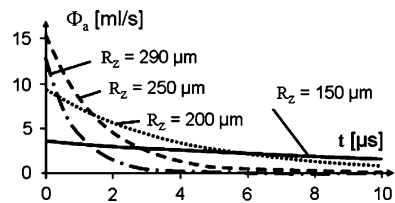
The flow  $\Phi_a$  through the outlet valve calculated with (404) and (405) is shown in Fig. 177. The same parameters were used as for Fig. 175. The comparison of these two figures shows that the volume flow through the outlet valve is rising with the radius of the valve seat and the feed channels. However, if the dead volume of the valves is taken into account, there is an optimum radius. If this optimum radius is exceeded, the flow through the outlet valve becomes smaller again, although the time which is required to empty the pump chamber gets shorter. The reason for this is that the pressure difference which is available to push the fluid through the outlet valve is smaller for a larger dead volume and the volume which can be delivered by one pump stroke is reduced also. This is recognized in Fig. 176: the position of the point A is moving to the right when the dead volume of the inlet valve becomes larger.

On the other hand, for small radii of feed channels and valve seats, the volume flow  $\Phi_a$  through the outlet valve is rising with increasing radii because the flow resistance  $R_{fl}$  is decreasing.

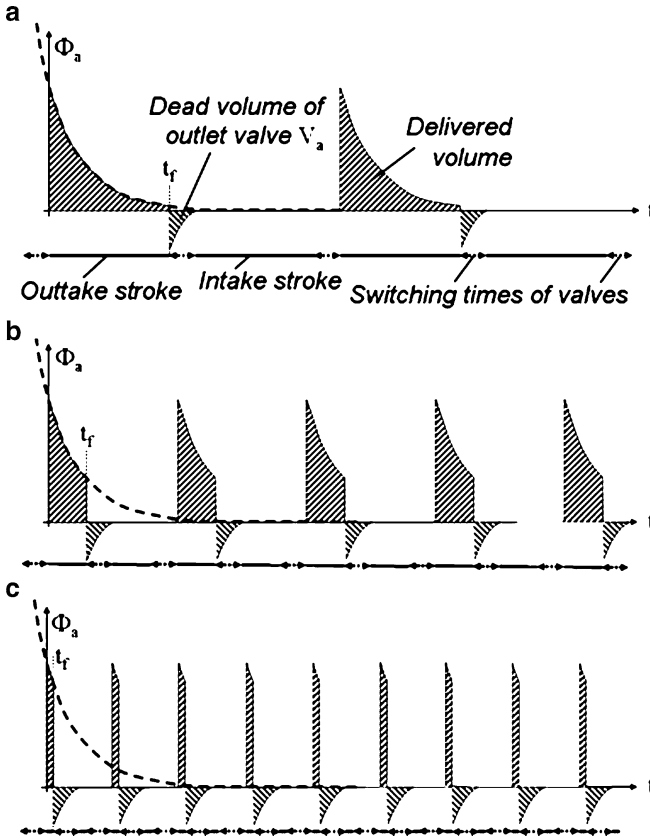
The flow rate of a reciprocating pump consists of repeated strokes of fluid volumes leaving the outlet. This is shown in Fig. 178 for different pumping frequencies. At a low frequency (Fig. 178a), nearly all of the volume in the actuator is delivered before the actuator is switched to its lower characteristic curve at the time  $t_f$ . Then the outlet valve is closing and its dead volume is flowing in backward direction during the switching time of the valve.

Next the intake stroke starts, the inlet valve opens, and fluid is sucked into the pump chamber. The actuator is switched to its upper characteristic curve and the dead volume of the inlet valve is flowing in backward direction. No flow is observed at the outlet during this time.

**Fig. 177** Flow through the outlet valve of a micropump as a function of time and the radius of the feed channel calculated with (404) and for the dead volumes of the valves (405)







**Fig. 178** Flow  $\Phi_a$  through the outlet valve of a pump as a function of time for (a) a low, (b) a medium, and (c) a high frequency of the pump actuator. The dashed area represents the delivered volume

A new pump cycle starts when the outlet valve opens again and fluid is pressed out of the pump chamber.

When the pump frequency is low (cf. Fig. 178a), most of the time there is only a small or even no volume flow at the outlet of the pump. At a somewhat higher frequency (Fig. 178b), only the part with a large flow through the outlet valve is used and the part with a smaller flow rate is forgone to start the next pump cycle earlier. As a consequence, the mean volume flow generated by the pump (the integral over  $\Phi_a$  divided by the time of a pump cycle) is larger than at a lower frequency.

If the frequency is enhanced even more, the volume delivered in a pump cycle in forward direction is not much more than flowing backward, and the actuator is only moving the dead volume of the valves back and forth. As a result, there is an optimum frequency at which the pump delivers the maximum possible flow.

The volume  $V_F$  delivered by one pump stroke is the displaced volume  $\Delta V_A$  at the time  $t_F$  when the actuator is switched from the upper characteristic curve to the

lower one ( $B$  in Fig. 176) minus the dead volume  $V_a$  of the outlet valve. This volume  $V_F$  delivered by one pump stroke needs to be multiplied with the pump frequency  $f_P$  to obtain the flow rate  $\Phi_P$  of the pump. With (403) it is obtained:

$$\Phi_P = f_P V_F = f_P [\Delta V_A(t_F) - \Delta V_A(0) - V_a] \quad (406)$$

$$\Rightarrow \Phi_P = f_P \left\{ \Delta V_{\max} \left( 1 - \frac{\Delta p_P}{\Delta p_{\max}} - \frac{V_e}{\Delta V_{\max}} \right) \left[ 1 - e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t_f}{R_{fl}}} \right] - V_a \right\} \quad (407)$$

The time  $t_f$  when the actuator is switched from one characteristic curve to the other can be estimated from the switching time of the valves and the frequency of the pump actuator. If it is assumed that the switching time  $t_v$  of inlet and outlet valve are the same and the duration of intake and outtake stroke are equal, it is obtained:

$$t_f = \frac{1}{2} \frac{1}{f_P} - t_v \quad (408)$$

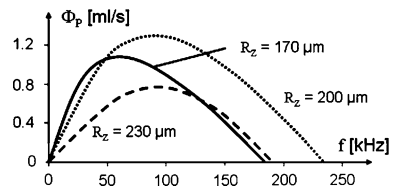
The switching times  $t_v$  of inlet and outlet valves are estimated by (390) on page 223. The dead volume of the valves is taken from (405) and the flow resistance  $R_{fl}$  of the valves was calculated with (355) (page 207) for a passive valve with circular cross-section and feed channels and a large stroke [the first term in the large parenthesis in the denominator of (355) is neglected]. The pressure difference over the valves during the back flow was estimated by the maximum pressure  $\Delta p_{\max}$  which the actuator is able to generate:

$$R_{fl,Z} = 8 \eta \frac{4 L_Z}{D_{fl,Z}^2 A_Z} = \frac{8 \eta L_Z}{\pi R_Z^4} \quad \text{and} \quad t_v = \frac{R_{fl,Z} V_V}{\Delta p_V} = \frac{4 \eta L_Z}{R_Z \Delta p_{\max}} \quad (409)$$

By inserting (408) and (409) into (407), the flow rate of a pump was calculated as a function of the driving frequency, and the radii of the feed channels and the valve seats. For the other parameters, the same values as for the calculation shown in Fig. 175 were used. In Fig. 179, the result of this calculation is shown.

For small frequencies, the flow rate is increasing proportional to the frequency because the volume displaced by each pump stroke is delivered more often. At higher frequencies, the volume  $V_F$  delivered by each pump stroke is decreasing because the actuator is switched to the lower characteristic curve before all the volume which could be displaced by the actuator has left the outlet valve and because  $V_F$  is no longer large compared with the dead volume of the valves. However, the flow rate is

**Fig. 179** Flow  $\Phi_P$  of a micropump as a function of actuator frequency  $f$  and the radius  $R_Z$  of feed channels and valve seats as calculated with (407)



decreasing with increasing frequency at very high frequencies (The scale is in kHz.). It is virtually impossible for a pump actuator to generate such high frequencies. In the literature, there was reported on a micropump which started to pump in reverse direction when the resonance frequency of the moving fluid and the valves was approached [64]. Here the effect of the acceleration of the fluid and the valve bodies and resonance effects have not been taken into account. Such effects occur only at large flow velocities which rarely are found at micropumps.

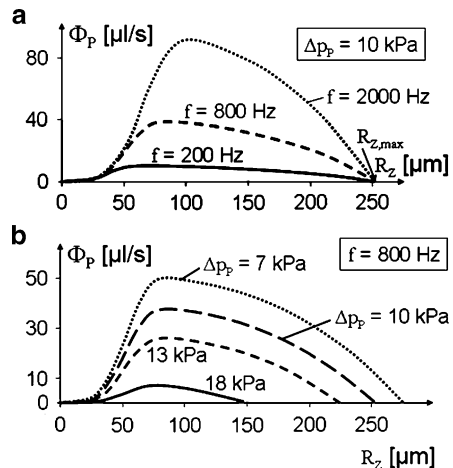
It is also seen in Fig. 179 that the slope of the curve at small frequencies is a function of the radius of feed channels and the valve seats. The effect of the size of the valves and feed channels is seen better when the pump flow calculated with (407) is drawn as a function of the radius of these components (cf. Fig. 180).

In Fig. 180a, the pump flow is shown for different actuator frequencies. It is clearly seen that the radius  $R_Z$  at which the largest flow is achieved is a function of frequency. If the radius is a bit larger than the optimum value, not much of the possible flow is lost. However, if the size of the valves and feed channels is only a bit too small, this may cause a remarkable loss in flow.

It is also seen in that figure that the maximum radius  $R_{Z,max}$  at which a flow can be generated is not a function of frequency. Figure 180b shows that  $R_{Z,max}$  is a function of the counter pressure  $\Delta p_p$ . For a given radius of the feed channels, there is maximum counter pressure  $\Delta p_{p,m}$  which limits the range of operation of the pump. At  $R_{Z,max}$ , the flow resistance  $R_{fl}$  of the feed channels and the valves is very small and the exponential function in (407) can be neglected compared with 1. At  $R_{Z,max}$ , the flow calculated with (407) is zero:

$$0 = f_P \left\{ \Delta V_{max} \left( 1 - \frac{\Delta p_p}{\Delta p_{max}} \right) - V_e - V_a \right\} \Rightarrow \Delta p_{p,m} = \left( 1 - \frac{V_e + V_a}{\Delta V_{max}} \right) \Delta p_{max}. \quad (410)$$

The above equation means that the maximum pressure  $\Delta p_p$  which can be generated by a pump is only a function of the maximum pressure  $\Delta p_{max}$  generated by the actuator, the ratio of the sum of the dead volumes of inlet  $V_e$  and outlet  $V_a$  valve



**Fig. 180** Flow  $\Phi_P$  of a micropump as a function of the radius  $R_Z$  of feed channels and valve seats and (a) the actuator frequency  $f$  and (b) the counter pressure  $\Delta p_p$ , respectively, calculated with (407)

and the maximum volume  $\Delta V_{\max}$  which can be displaced by the actuator. The maximum pressure of a pump is a function of the size of the valves, because their dead volume is a function of that size. If (405) is employed to calculate the dead volume, the maximum pressure  $\Delta p_{P,m}$  of a pump as a function of the radius  $R_Z$  of the valve seat is:

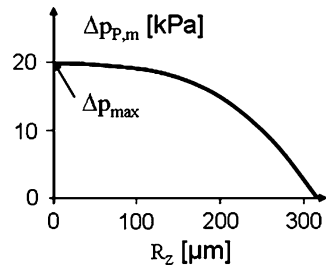
$$\Delta p_{P,m} = \left( 1 - \frac{\pi R_Z^3}{\Delta V_{\max}} \right) \Delta p_{\max} \tag{411}$$

The maximum pressure  $\Delta p_{P,m}$  generated by the pump with the parameters used above was calculated with the above equation and drawn as a function of  $R_Z$  in Fig. 181. It is clearly seen that the largest pressure can be generated with a pump which has very small valves. The smaller the valves are the more the maximum pressure  $\Delta p_{P,m}$  generated by the pump approaches the maximum pressure  $\Delta p_{\max}$  generated by the actuator. However, small valves will result in a small flow (cf. Fig. 180). Therefore, a compromise needs to be made between a large flow and a large pressure generated.

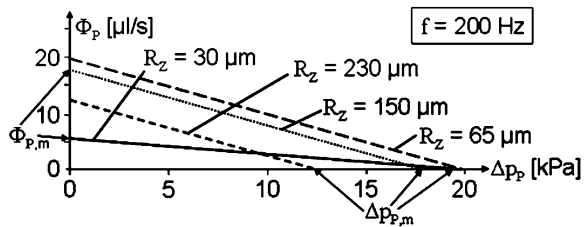
The flow generated by a pump as a function of the counter pressure is called the characteristic curve of the pump. This curve is obtained when (407) is plotted as a function of the counter pressure. Figure 182 shows the characteristic curve of a pump with the parameters used above for different radii of feed channels and valve seats. The characteristic curve is a straight line between the maximum flow  $\Phi_{P,m}$  and the maximum pressure  $\Delta p_{P,m}$  which can be generated by the pump.

Obviously, the flow is rising for small valves with their size while the maximum pressure which can be achieved is nearly not changing (see also Fig. 181). In this example, the optimum radius is reached at approximately 65  $\mu\text{m}$ . When the radius  $R_Z$  exceeds this value, the characteristic curve is shifting left and down nearly

**Fig. 181** Maximum pressure  $\Delta p_{P,m}$  which can be generated with a pump as a function of the radius  $R_Z$  of feed channels and valve seats



**Fig. 182** Characteristic curve of a pump as a function of the radius  $R_Z$  of feed channels and valve seats



parallel. Thus, maximum flow rate and maximum pressure both are reducing when the radii  $R_z$  of the feed channels and the valves are increasing.

If the dead volume and the flow resistance of the valves would have no effect on maximum flow rate and maximum pressure  $\Delta p_{P,m}$  which can be generated by the pump, the maximum flow  $\Phi_{P,m}$  would be equal to the frequency times the maximum pressure which can be generated by the actuator (20  $\mu\text{L/s}$  in Fig. 182) and the maximum pressure generated by the pump would be equal to the maximum pressure  $\Delta p_{\text{max}}$  of the actuator (20 kPa in Fig. 182). It is seen in the figure that at the optimum radius these values are nearly reached (19.8  $\mu\text{L/s}$  and 19.8 kPa, respectively). However, even larger flow rates can be achieved, if active valves are employed and the flow resistance of the valve is reduced by designing them larger. The disadvantage of this concept is that three actuators and a suitable electronic control are required. An additional advantage is that a pump with active valves can pump also in reverse direction.

The *fluidic power of a pump*  $P_P$  is the product of the flow rate  $\Phi_P$  and the counter pressure  $\Delta p_P$ . Thus, the fluidic power is calculated as the product of (407) with the counter pressure. Since the characteristic curve is a linear function of the pressure, the result of multiplying with the pressure is a parabola and the maximum  $P_{P,m}$  of that parabola is a fourth of the product of maximum counter pressure  $\Delta p_{P,m}$  and maximum flow rate  $\Phi_{P,m}$ :

$$P_{P,\text{max}} = \frac{1}{4} \Phi_{P,m} \Delta p_{P,m} \tag{412}$$

The fluidic power for the graphs shown in Fig. 182 is drawn in Fig. 183. The input power of the piezo actuator is not much affected by the counter pressure and the flow generated (cf. page 150). Therefore, the power output shown in Fig. 183 is nearly proportional to the efficiency of the pump. If micropumps are to be designed for portable or implantable devices which are powered by batteries, this may be an important issue. Since the fluidic power  $P_P$  is a parabolic function of the counter pressure, fluidic power and efficiency of a pump are maximum at half of the maximum counter pressure  $P_{P,m}$ . Therefore, when energy consumption of a pump is a concern, it should be designed such that the typical counter pressure is half of the maximum counter pressure.

When the radii  $R_z$  of the feed channels and the valve seats are increased, the maximum of the fluidic power is increasing, while the optimum counter pressure is nearly not decreasing. When the maximum fluidic power which can be achieved by

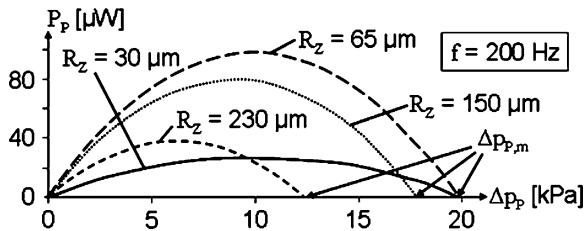


Fig. 183 Fluidic power of a micropump as a function of the counter pressure and the radii of feed channels and the valve seats

increasing the radii  $R_Z$  is exceeded, the optimum counter pressure and the maximum fluidic power are decreasing.

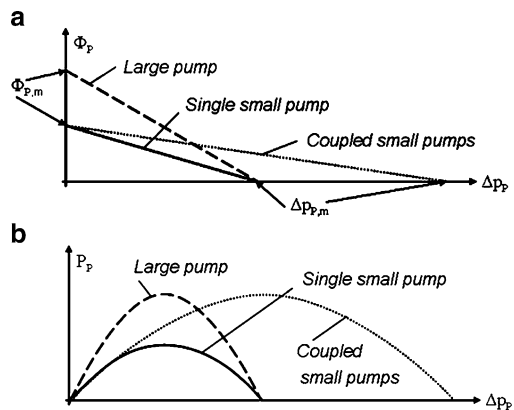
When a pump is employed for a dosing application, it is desirable that the flow rate is not or only a weak function of the counter pressure. This can be achieved by limiting the maximum stroke of the actuator (designing a very shallow pump chamber) to much less than the maximum stroke which the actuator could perform without that restriction. As a consequence, the pressure generated by the actuator during the stroke is nearly constant.

Even better is employing an actuator for which the force generated is not a function of the stroke such as the lateral capacitive force (cf. page 132), evaporation of a liquid in a closed volume (cf. page 171), and an external pneumatic or hydraulic drive.

Another option is to design two small pumps working in series instead of a larger pump with the same volume. The first of the two pumps needs to be driven such that its outtake stroke is over before the intake of the second pump starts. If both pumps are driven in phase, they behave similar as a single pump with a larger pump chamber.

If the two pumps are not driven in phase, each of them needs to overcome a smaller counter pressure and they have a smaller maximum flow rate, because their pump chambers are smaller than the chamber of a single pump with the same volume as both pumps together. The resulting characteristic curves are shown in Fig. 184a. If the maximum volume  $V_{max}$  displaced by the actuator and the dead volumes of inlet  $V_e$  and outlet  $V_a$  valves are all reduced by the same factor, e.g. 2, the maximum flow  $\Phi_{p,m}$  is also reduced by that factor (cf. 407) and the maximum pressure generated by such a small pump is not changing (cf. 410). If two pumps are working in series and are driven with a phase shift, the maximum pressure generated by both pumps together is doubled. As shown in Fig. 184a, the slope of the characteristic curve of the coupled pumps is significantly reduced compared with the larger single pump and a change of the counter pressure  $\Delta p_p$  has a much smaller effect on the flow rate  $\Phi_p$ . The flow rate can be adjusted then by choosing the driving frequency of the actuator.

Another interesting effect of coupled pumps driven not in phase is that the maximum efficiency of pumping is shifted to a larger counter pressure. This is



**Fig. 184** (a) Characteristic curve of a pump, a small pump, and two small pumps in series driven with a phase shift and (b) fluidic power of these pumps

shown in Fig. 184b. Since the maximum fluidic power is a fourth of the product of maximum flow  $\Phi_{P,m}$  and maximum counter pressure  $\Delta p_{P,m}$  (cf. 412), the maximum fluidic power of two small coupled pumps with half of the maximum flow is the same as for a larger pump. But the maximum is at a larger counter pressure and a certain change of the counter pressure does not affect the fluidic power so much. Therefore, coupling two or even more pumps may save battery power in certain applications.

Leaking valves show a similar effect as larger dead volumes. Thus, in all equations such as (407) (page 236) and (410), the volume lost by leaking needs to be added to the dead volume of the inlet  $V_e$  or outlet  $V_a$  valve, and both the flow rate and the pressure which can be generated by the pump are decreasing.

When a liquid is pumped and a *gas bubble is in the pump chamber*, the volume of the bubble is decreased and increased by the pressure change in the pump. This volume changes  $\Delta V_B$  do not contribute to the flow rate of the pump and need to be subtracted from the volume  $V_F$  which is pumped with each pump cycle. The volume change of the bubble has a similar effect as the dead volumes of inlet  $V_e$  and outlet  $V_a$  valve (cf. Fig. 176 on page 233). When the pressure in the pump chamber is increased by the actuator, the dead volume of the inlet valve is displaced out of the pump and the volume of the bubble is decreased by  $\Delta V_B$ , and when the pressure in the pump is decreased again, the dead volume of the outlet valve is flowing in backward direction and the volume of the bubble is increased before new liquid is taken into the pump. Therefore, the effect of a bubble is similar to the one of larger dead volumes of inlet and outlet valve.

The volume change  $\Delta V_B$  of the bubble in the pump chamber can be calculated with the ideal gas law [(295) on page 167] which means that the product of pressure and volume of the bubble is constant (as long as the temperature does not change). When the inlet valve opens, the pressure and volume of the bubble are denoted by  $p_e$  and  $V_{B,e}$ , respectively. When the outlet valve opens and when there is a pressure of 101.3 kPa, they are denoted by  $p_a$ ,  $V_{B,a}$ , and by  $p_0$ ,  $V_{B,0}$ , respectively:

$$p_e V_{B,e} = p_a V_{B,a} = p_0 V_{B,0} \tag{413}$$

$$\Rightarrow \Delta V_B = V_{B,e} - V_{B,a} = \left( \frac{1}{p_e} - \frac{1}{p_a} \right) p_0 V_{B,0} = \frac{\Delta p_P}{p_e p_a} p_0 V_{B,0}. \tag{414}$$

Inserting this into (407) (page 236) as an additional dead volume of inlet and outlet valve yields:

$$\Phi_P = f_P \left\{ \left[ \Delta V_{\max} \left( 1 - \frac{\Delta p_P}{\Delta p_{\max}} \right) - V_e - \Delta V_B \right] \left[ 1 - e^{-\frac{\Delta p_{\max}}{\Delta V_{\max}} \frac{t_f}{R_{fl}}} \right] - V_a - \Delta V_B \right\} \tag{415}$$

As a consequence of the above equation, the pump is no longer able to pump any liquid when the bubble becomes too large. Thus, besides the problem that bubbles can block the valves because the gap over the valve seat is narrower than the feed channels

(cf. page 118), the volume change of bubbles in the pump chamber decreases or even stops the flow. Special microstructures before a micropump can separate bubbles from the liquid flow to avoid such problems with bubbles (cf. page 118).

The volume of a bubble is limited by the dead volume  $V_P$  of the pump which is defined as the inner volume of the pump between the valve seats of inlet and outlet valve. The flow rate calculated with (415) can be larger than zero when the term in the curly braces is larger than zero. The exponential function can be assumed to be zero because the time of the delivering pulse of the pump will be made large to achieve a larger flow rate:

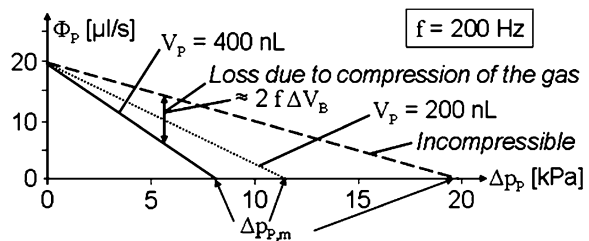
$$\Delta V_{\max} \left( 1 - \frac{\Delta p_P}{\Delta p_{\max}} \right) \geq V_e + V_a + 2 \frac{\Delta p_P}{p_e p_a} p_0 V_P \tag{416}$$

$$\Rightarrow V_P \leq \left[ \Delta V_{\max} \left( 1 - \frac{\Delta p_P}{\Delta p_{\max}} \right) - V_e - V_a \right] \frac{p_e p_a}{2 p_0 \Delta p_P} \tag{417}$$

If the dead volume  $V_P$  of a pump is small enough to fulfill the above equation, the pump is also able to pump a gas. As a consequence, a pump with such a small dead volume is also self-priming, i.e., the pump can pump the air out of a pipe or hose connected to its inlet, and, this way, suck in a liquid. Besides this, it has to be concluded that the dead volume of a pump should be as small as possible to achieve a large flow rate and especially a large pressure.

According to (414), the volume change  $\Delta V_B$  of a bubble in the pump chamber is proportional to the pressure difference over the pump  $\Delta p_P$ . That is, a bubble has only little effect on the flow rate when the pressure difference over the pump is small. The larger the counter pressure is, the more the flow rate is decreased by the volume change of the pump. In Fig. 185, this is shown for the case that the bubble is as large as the dead volume of the pump, i.e., only gas is pumped. The same parameters are assumed as for Fig. 182 and a radius  $R_Z$  of 65  $\mu\text{m}$ . The characteristic curves of the pump pumping an incompressible fluid is shown together with the curves of a compressible gas and a dead volume  $V_P$  of the pump which is a factor of 2 (200 nL) and 4 (400 nL) larger than the maximum displacement  $\Delta V_{\max}$  of the actuator, respectively.

The maximum counter pressure  $\Delta p_{P,m}$  which can be generated by a pump pumping a compressible fluid is found when the flow rate calculated with (415) is zero. This is the case when the term in the curly braces is zero. The exponential



**Fig. 185** Characteristic curve of a micropump pumping an incompressible fluid and a compressible gas with (theoretically) the same viscosity



function can be assumed to be zero because the time of the delivering pulse of the pump will be made large to achieve a larger counter pressure:

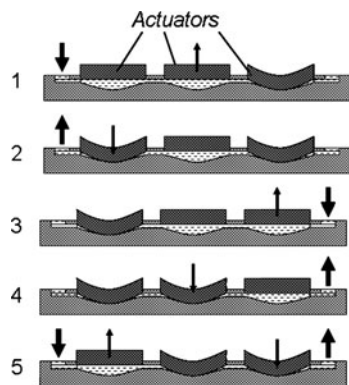
$$\Delta p_{p,m} = \frac{\Delta V_{\max} - V_e - V_a}{2 \frac{p_o V_p}{p_e p_a} + \frac{\Delta V_{\max}}{\Delta p_{\max}}} \tag{418}$$

As shown above, both the maximum counter pressure and the maximum flow rate are reduced by the need to open or close the valves and to pump their dead volumes. This can be avoided when a *pump with active valves* is employed. As a result, the pump cycle is as shown in Fig. 174 instead of Fig. 176 (on page 230 and 233, respectively). The flow through the outlet valve is calculated with (401) and the flow passes through the outlet the quicker, the larger the radii of the feed channels and the valve are (cf. Fig. 175, page 232). Thus, a pump with active valves achieves a larger maximum flow rate and a larger maximum counter pressure. The reason for this is that the energy of the pump actuator is only used for pumping the fluid and not to open and close the valves. However, this energy is provided now by the actuators of the valves, and a similar result could be obtained with a somewhat larger pump actuator and passive valves.

A clear advantage of a pump with active valves is that the pump direction can be reversed by controlling valve switching. Therefore, a pump with active valves may replace two pumps and additional valves in certain applications. A special case of a pump with active valves is a peristaltic pump, which employs three or more equal pump chambers driven in consecutive order (cf. Fig. 186).

A pump delivering air which shows a dead volume of the valves as large as the pump chamber is shown in Fig. 187 [65]. In a pump chamber which is formed by two bell-shaped walls, there are mounted two membranes. In the membranes, there are orifices staggered to each other such that no air can pass through when the membranes are in touch to each other.

Therefore, the pump is sealed when both membranes are at the bottom of the pump chamber (Fig. 187a). Both membranes are insulated to the surrounding and contain electrodes. The membranes are held down by applying a voltage between the upper membrane and the bottom of the pump chamber. Then, voltage is applied between the

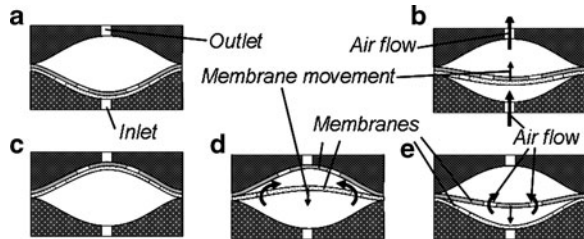


**Fig. 186** Pump cycle of a peristaltic pump

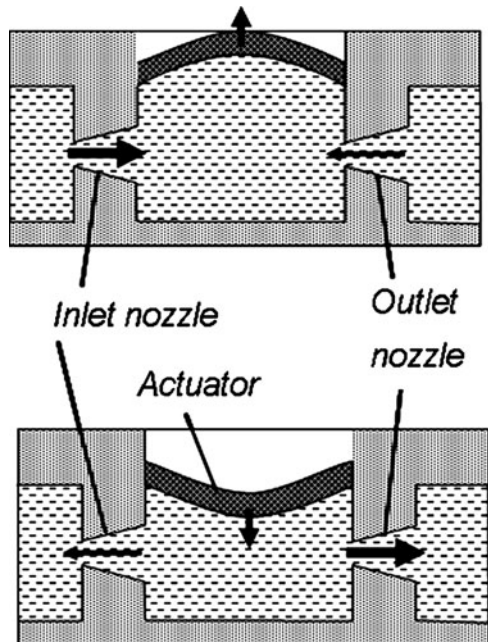
lower membrane and the upper part of the pump chamber. This way, both membranes are moving up and delivering air to the outlet of the pump and taking in fresh air through the inlet (Fig. 187b). Since electrostatic forces are strong only at small distances, the bell-shaped walls facilitate moving of the membranes. The movement starts at the rim of the membranes where there is only a small distance to the upper wall and then continues until the entire membranes are in touch to the upper wall (Fig. 187c).

Then only the lower membrane is attracted down to the bottom of the pump chamber and no air is entering or leaving the pump, because the outlet is sealed by the upper membrane (Fig. 187d). During this membrane movement, the flow resistance of the orifices of the membrane needs to be overcome. The upper membrane follows the lower one in the next step and again no air enters or leaves the pump because the lower membrane seals the inlet (Fig. 187e).

Obviously, the air volume in the pump chamber in every pump cycle needs to pass through the membranes two times, and, therefore, is the dead volume of the “valves”.



**Fig. 187** Cycle of a pump with two electrostatically driven membranes [65]. © [1997] IEEE



**Fig. 188** Valveless micropump

For certain applications, it may be an advantage of this pump that it can be driven in backward direction.

It is also possible to design valves without any moving mechanical parts. Pumps employing such valves are sometimes called “valveless” pumps. Their valves consist of nozzles which show a larger flow resistance in one direction than in the opposite one (see Fig. 188). The back flow in the undesired direction is compensated by a larger flow in the desired one. However, such valves have a huge dead volume, and, as a consequence, they are not able to work against large counter pressures. Besides this, they are open in both directions when the actuator is not running and they achieve only a small efficiency. It is an advantage of such pumps that their valves are very rugged and easy to fabricate. However, other kinds of valves are also rugged enough.

The pumps described so far are all reciprocating pumps which generate a more or less pulsing flow. For some applications, this is not desirable. Solutions for that problem are *aperiodic pumps*. One aperiodic pumping principle has already been described earlier in this book: the electro-osmotic micropump (cf. page 122). Another possibility is to design a small chamber and fill it with a sorption agent which draws in a liquid continuously until it is completely filled [66].

Another interesting design is the electrohydrodynamic pump (EHD) [67]. An EHD consists of two gratings which spread across a channel. Between the gratings there is a certain distance. One grating is fabricated such that it shows a sharp edge pointing to the other grating. When the negative pole of a voltage is applied to the grating with the sharp edge and the positive to the other one, electrons are injected from the sharp edges into the liquid and they are accelerated towards the grating with the positive pole. Due to friction in the liquid not only the electrons but also the entire liquid is accelerated.

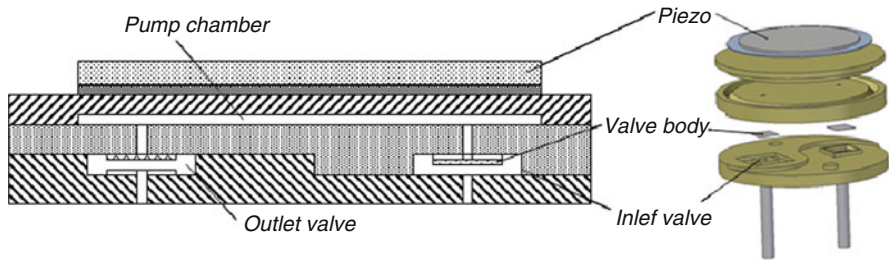
The EHD principle only works if the liquid is not conductive. Therefore, water cannot be delivered but certain oils. Large flows are achieved but only a very small pressure difference can be overcome.

## Exercises

### Problem 38

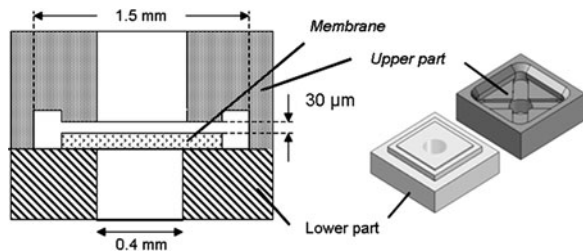
Figure E29 shows the cross-section of a micropump [68]. The properties of the pump are listed in the following table.

Diameter of pump chamber and actuator	12 mm
Deflection of the piezo actuator	20 $\mu\text{m}$
Length of the edge of the quadratic valves	1.5 mm
Length of the valve gap in the direction of the flow	500 $\mu\text{m}$
Stroke of the valve	30 $\mu\text{m}$
Diameter of the feed channels of the valve	0.4 mm
Length of the feed channels of each valve	1 mm
Viscosity of water	1 mPa s
Maximum pressure difference $\Delta p_{\text{max}}$ generated by the actuator	100 kPa



**Fig. E29** Micropump with a piezo actuator and polymer plates as valve bodies. (Reprinted from [68] with permission from Elsevier)

**Fig. E30** Valve of the micropump shown in Fig. E29. (Reprinted from [68] with permission from Elsevier)



- (a) Please calculate the maximum volume change generated by the actuator.
- (b) Please calculate the length of the pump pulse. At what frequency does the micropump achieve the maximum flow rate if the switching times of the valves are negligibly short and the maximum volume change generated by the actuator is  $1 \mu\text{L}$ ?
- (c) In the valves of the pump, freely movable square plates from polymer are installed as valve bodies as shown in Fig. E30. The length of the edges of these plates is  $1.4 \text{ mm}$ . Please calculate the dead volume of the passive valves. Neglect the volume of the grooves in the upper part.
- (d) What is the switching time of the passive valve according to (390) (page 223), if water is used as the fluid? Assume that the length of the feed channels and the pressure drop over them are  $1 \text{ mm}$  and  $50 \text{ kPa}$ , respectively.
- (e) How could the design be improved?

**Problem 39**

It is your objective to design flap valves for a micropump. Assume that the mass of the fluid and the valve body can be neglected. As an approximation, assume that the pump is driven at its optimum frequency and that the stroke of the valves is large

also at the beginning of the pump pulses. The following data are available for the calculation.

Maximum volume change generated by the actuator $\Delta V_{\max}$	600 nL
Maximum pressure difference generated by the actuator $\Delta p_{\max}$	24 kPa
Pressure difference between inlet and outlet $p_a - p_e$	12 kPa
Length of the feed channels to the valves	400 $\mu\text{m}$
Viscosity of water	1 mPa s
Length of the valve beam	1 mm
Young's modulus of the beam	120 GPa
Thickness of the beam	10 $\mu\text{m}$

- Calculate the radius of feed channels and inlet and outlet (assumed to be equal) required for the maximum flow rate of the pump.
- What is the maximum flow rate that can be achieved?
- With which tolerance needs the radius of the valve orifice  $R_Z$  to be manufactured, if the flow rate may deviate from the maximum flow rate calculated in (b) not more than 0.1 mL/s? Assume for this calculation that all other dimensions are exactly kept to the desired measures. (Hint: For this problem, it is advantageous to employ a spreadsheet program.)

### ***Problem 40***

For this problem, the data of the following table are given:

Radius of the glass sphere	200 $\mu\text{m}$
Density of glass	2,600 kg/m <sup>3</sup>
Radius of the feed channels	100 $\mu\text{m}$
Acceleration of gravity	9.81 m/s <sup>2</sup>
Length of the feed channels	1 mm
Length of the valve gap	200 $\mu\text{m}$
Young's modulus of silicon	130 GPa
Width of the beam	600 $\mu\text{m}$
Thickness of the beam	20 $\mu\text{m}$
Viscosity of water	1 mPa s
Maximum pressure difference generated by the actuator $\Delta p_{\max}$	200 kPa
Maximum volume change generated by the actuator $\Delta V_{\max}$	1 $\mu\text{L}$

Derive the equation for the flow of a micropump without counter pressure, if instead of a flap valve a freely moving sphere is employed in the passive valves. Please follow the subsequent questions:

- If the sphere is made of glass with a diameter of 400  $\mu\text{m}$ , what is the minimum pressure difference required to lift the sphere in the gravitational field and to open the valve? Please neglect the buoyant force of the sphere in the fluid and any stiction.

- (b) How large should be the stroke of the valve (Rule of thumb)?
- (c) The stroke as a function of the pressure drop for a sphere cannot be found easily because it is a function of the flow situation inside of the valve. However, it appears to be reasonable to assume that the sphere is moved up to a stop 55  $\mu\text{m}$  above the valve seat when the pressure difference is 60 Pa.  
Please calculate how long a 600- $\mu\text{m}$  wide and 20- $\mu\text{m}$  thick beam from silicon needs to be which is employed instead of the sphere in (a). The beam shall show a stroke of 55  $\mu\text{m}$  at a pressure difference of 60 Pa.
- (d) What is the dead volume of the microvalve (stroke = 55  $\mu\text{m}$ )?
- (e) Calculate the switching time of the valve, if the effective pressure difference over the feed channels is 80 Pa.
- (f) Please calculate the pump pulse time on the one hand with the approximation of widely opened valves and on the other hand with the effect of the valve gap.
- (g) Why is it advantageous to limit the stroke of the valves (both for a sphere and a beam) by a stop above the valve seat?
- (h) What is the maximum pump frequency, if the switching time of the valves is taken into account and the pump pulse time is 150  $\mu\text{s}$ ? (The question whether the actuator is able to generate this frequency is not discussed here.)
- (i) What is the delivered flow volume per pump pulse without counter pressure?
- (j) What is the maximum flow rate of the pump without counter pressure?

### Problem 41

For a diagnostic device, a liquid flow ( $\eta = 1 \text{ mPa s}$ ) of 1.5  $\mu\text{L/s}$  shall be delivered against a pressure difference of 50 kPa. You have micropumps with flap valves on stock which fulfill the following specifications:

Radius of the feed channels	24 $\mu\text{m}$
Length of the feed channels	0.5 mm
Length of the valve gap	200 $\mu\text{m}$
Young's modulus of the beam	0.6 GPa
Length of the beam	1 mm
Thickness of the beam	15 $\mu\text{m}$
Maximum pressure difference generated by the actuator $\Delta p_{\text{max}}$	100 kPa
Maximum volume change generated by the actuator $\Delta V_{\text{max}}$	50 $\mu\text{L}$

- (a) Can the objective be fulfilled by the micropump?
- (b) The flow rate at a counter pressure of 50 kPa now shall be enhanced. To achieve this, you first try to change the size of the feed channels. What radius should the feed channels obtain to achieve the maximum possible flow rate at the given counter pressure? How large is the maximum flow rate? How much can the flow rate be enhanced by changing the diameter of the feed channels? (A spreadsheet program or a programmable pocket calculator will help to solve this problem.)

- (c) Now prove the alternative that two down sized micropumps driven in series are employed, if each of them achieves half of the flow rate calculated in (a) but the same maximum pressure. What is in this case the flow rate at 50 kPa counter pressure?
- (d) Calculate the fluidic power of the pump in (b) at a counter pressure of 50 kPa.

**Problem 42**

For the following problem, the data of the following table are given:

Diameter of the pump chamber	4 mm
Minimum flow rate	100 $\mu$ L/h
Maximum counter pressure	20 kPa
Piezo electric modulus $d_{31}$	$-2.14 \times 10^{-10}$ m/V
Thickness of the piezo layer	100 $\mu$ m
Absolute permittivity	$8.9 \times 10^{-12}$ A s/(V m)
Relative permittivity of air	1
Young’s modulus of piezo ceramic	66.6 GPa
Young’s modulus of silicon dioxide	73 GPa
Gas constant $R_G$	8.314 J/(mol K)

For an insulin pump, an actuator needs to be chosen. An insulin pump can be mounted outside of the skin and deliver through a catheter into the body. As an alternative, the pump can be implanted into the body. Figure E31 shows an example of an implantable micropump.

- (a) Please calculate the voltage required to achieve the pressure difference with the piezo actuator. The piezo has the same diameter as the pump chamber and is glued onto a carrier from silicon dioxide, 50  $\mu$ m in thickness. The thickness of the glue can be neglected.
- (b) The voltage calculated in (a) is just enough to overcome the counter pressure but is not sufficient to deliver against this pressure. Therefore, you decide to apply 128 V. Please calculate the frequency at which the pump needs to run to achieve the specified flow rate at a counter pressure of 20 kPa. Do not take into account the dead volume of the valves.



**Fig. E31** Implantable micropump [69] (Courtesy of Debiotech SA/Switzerland)

- (c) Please calculate the minimum voltage required to achieve the specified pressure difference by an electrostatic actuator. Neglect the elastic force of the membrane and assume that there is air between the electrodes. The distance between membrane and counter electrode is  $10\ \mu\text{m}$ .
- (d) If for an implantable pump for safety reasons there is allowed only a voltage smaller than  $36\ \text{V}$ , how needs the design to be changed to achieve the required counter pressure.
- (e) How much would a thermo-pneumatic actuator need to be heated up to achieve the required pressure difference?
- (f) For some applications, a micropump driven hydraulically by the blood pressure could be of interest. It could, e.g., be self-controlling when some drug against high blood pressure is delivered more when the blood pressure or the pulse beat is high. How large would the flow rate of such a pump be if its displaced volume per pulse is  $0.02\ \mu\text{L}$  at a typical heart frequency of  $1\ \text{Hz}$ ?



# Microdosing

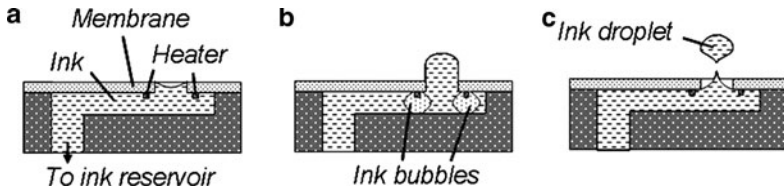
Dosing of small amounts of liquids is a widespread application of microfluidics. The most common application is ink-jet printers, but dosing of adhesives and lubricants, and chemical and biological substances in pharmaceutical industry also occur often. Obviously, smaller amounts of a liquid can be dosed precisely when miniature devices are used.

Microdosing can be achieved by employing a micropump with a well-defined stroke which is only a weak or even no function of the counter pressure (cf. page 240). A reservoir with a well-defined pressure and a microvalve opened for a well-defined time and a well-defined stroke can also be employed for microdosing. Less accuracy is required for a microvalve or a micropump when they are combined with a precise flow sensor (cf. page 289). However, the combination of an actuator with a flow sensor is more complex than a single device and the control loop needs to achieve the required precision also.

The most numerous microdosing systems are the heads of ink-jet printers. Such heads include a feed channel to the reservoir, a heater, and an orifice where the ink is ejected from (Fig. 189a). The heater consists of a conductor path with which an ink is evaporated generating a bubble in the liquid ink (Fig. 189b). The bubble then ejects an ink droplet through the orifice (Fig. 189c). After the droplet has been ejected and the bubbles condensate back into liquid ink, the feed channel is filled again by capillary force.

If the heater is supported by a membrane as shown in Fig. 189, the generated heat is not so much absorbed by the housing of the printer head but contributes more to the evaporating of ink and ejecting of the droplet. However, a printer head with the heater on the solid housing may be easier to fabricate and more rugged.

Droplets can also be ejected from a device with a piezo generating a quick pressure rise in a liquid column. Such devices are employed for dosing small amounts of, e.g., lubricants or adhesives applied to the bearings of, e.g., watches or small parts which are to be joined, respectively.



**Fig. 189** Ejection of a droplet from an ink-jet printer

# Analogies of Physical Domains

In microtechnique, often several basic components are combined to devices and several devices to systems. The basic elements employed in a system can belong to different physical domains such as mechanical, fluidic, electrical, and thermal. This chapter describes how to calculate the behavior of coupled elements and systems with the help of analogies.

The behavior of all physical systems is calculated from the laws of conservation of energy and momentum, or equilibrium of forces. As a consequence, different systems which show an analogy between their equations of energy and forces have the same analogy in the equations of motion, etc.

There are three types of energy: potential energy, kinetic energy, and energy lost by dissipation. Similarly, there are three types of forces: elastic forces, inertial forces, and frictional forces. As an example, the equations of energy and forces for a mass  $m_K$  fixed to the end of a spring with spring constant  $k$  (cf. Fig. 63 on page 89) damped with a damping constant  $D_R$  and of an electrical circuit with a resistance  $R_{el}$ , a coil with inductance  $L_{el}$ , and a capacitor with capacitance  $C_{el}$  in series (cf. Fig. 190) are as following:

The comparison of the above equations shows that they are analog to each other, if the following quantities are substituted by each other:

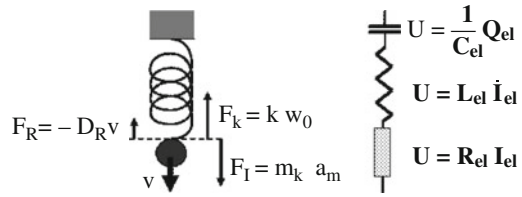
$$x \leftrightarrow Q_{el}, v \leftrightarrow I_{el}, a_m \leftrightarrow \dot{I}_{el}, k \leftrightarrow \frac{1}{C_{el}}, m_K \leftrightarrow L_{el}, D_R \leftrightarrow R_{el}.$$

Since all properties of both systems are calculated from their energy or force equations, the equation for a property of one system can be found from the solution of the other system just by the analogy.

For example, from (133) on page 89 the resonance frequency  $f_r = 2 \pi \omega_r$  of a mass fixed to a spring is known, and from the analogy above for an electrical circuit with a capacitor and a coil it is found:

$$f_r = \frac{1}{2\pi} \sqrt{\frac{k}{m_K}} \rightarrow f_r = \frac{1}{2\pi} \sqrt{\frac{1}{C_{el} L_{el}}}. \quad (419)$$

**Fig. 190** Mass fixed to a spring compared with capacitor, coil, and resistance in series



**Table 19** Equations of the energies and forces employed for the calculation of the behavior of a mass fixed to a spring and an electrical circuit consisting of a capacitor, a coil, and a resistance

	Mass fixed to a spring	Electrical circuit
Potential energy	$\frac{1}{2} k x^2$	$\frac{1}{2} \frac{Q_{el}^2}{C_{el}}$
Kinetic energy	$\frac{1}{2} m_K v^2$	$\frac{1}{2} L_{el} I_{el}^2$
Dissipated energy	$D_R v^2 t$	$R_{el} I_{el}^2 t$
Elastic force	$F_K = k x$	$U = \frac{Q_{el}}{C_{el}}$
Inertial force	$F_I = m_K a_m$	$U = L_{el} \dot{I}_{el}$
Frictional force	$F_R = D_R v$	$U = R_{el} I_{el}$

The influence of the electrical resistance on the resonance frequency of an electrical circuit may be found in a text book on electrotechnique and the influence of mechanical damping on the resonance frequency of a mass fixed to a spring can be derived from that:

$$f_r = \frac{1}{2\pi} \sqrt{\frac{1}{C_{el}L_{el}} - \frac{R_{el}^2}{4L_{el}^2}} \rightarrow f_r = \frac{1}{2\pi} \sqrt{\frac{k}{m_K} - \frac{D_R^2}{4m_K^2}} \tag{420}$$

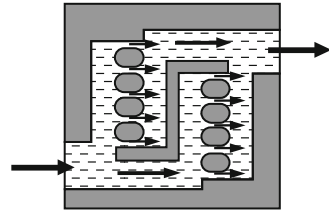
Such conclusions may be drawn only if the energy equations or equivalently the force equations really have the same form. For example, it would be wrong to derive from the above analogy the resonance frequency of a pendulum, because the gravitational force is not included in the set of equations above and has another form as the equations of the elastic force above, i.e., it is not proportional to \$x\$ or \$Q\_{el}\$, respectively:

$$F = m_K g_e \tag{421}$$

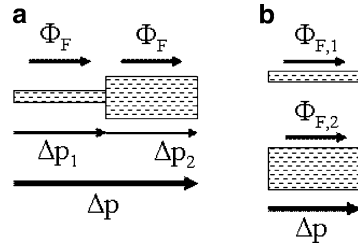
There is also an analogy to fluidic systems. The friction in a capillary is described by an equation [(178), page 108] which is similar to the frictional force in Table 19.

$$F_R = D_R v \leftrightarrow \Delta p = R_{fl} \Phi_F \Rightarrow F_R \leftrightarrow \Delta p, D_R \leftrightarrow R_{fl}, v \leftrightarrow \Phi_F \tag{422}$$

**Fig. 191** Flow through a system of microchannels



**Fig. 192** Pressure and volume flow at capillaries (a) in series and (b) parallel to each other



A very simple example is the combination of microchannels to a system of channels which, e.g., may serve as a filter for particles in a fluid. Figure 191 shows an example. The designer of such a filter needs to know, what is the throughput of the filter as a function of the pressure difference  $\Delta p$ . The flow  $\Phi_F$  through a single capillary with fluidic resistance  $R_{fl}$  is known from the Hagen Poiseuille equation (175) (page 107) and (178) (page 108), respectively.

The volume flow through two capillaries in series is the same in each capillary, because the fluid can neither disappear nor appear in between. The pressure drop over both capillaries obviously is the sum of the pressure differences over the individual ones (cf. Fig. 192a). With (422) it is obtained:

$$\Delta p = \Delta p_1 + \Delta p_2 = R_{fl,1}\Phi_F + R_{fl,2}\Phi_F = R_{fl}\Phi_F \Rightarrow R_{fl} = R_{fl,1} + R_{fl,2}. \quad (423)$$

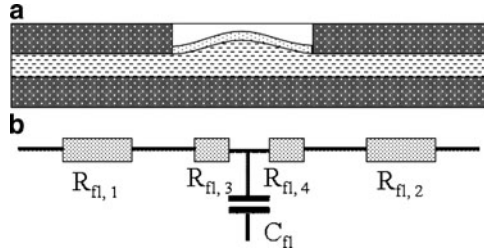
When two capillaries are arranged in parallel (Fig. 192b), the pressure difference is the same over both capillaries and the overall volume flow is the sum of the flows through the two capillaries. With (422) it is obtained:

$$\Phi_F = \Phi_{F,1} + \Phi_{F,2} = \frac{\Delta p}{R_{fl,1}} + \frac{\Delta p}{R_{fl,2}} = \frac{\Delta p}{R_{fl}} \Rightarrow \frac{1}{R_{fl}} = \frac{1}{R_{fl,1}} + \frac{1}{R_{fl,2}}. \quad (424)$$

Obviously, the flow resistances of channel systems are calculated in the same way as the electrical resistance of a system of resistances. Besides this, (422) is of the same kind as Ohm’s law, if the pressure difference is associated with the voltage and the volume flow with the electrical current.

To fulfill the analogy to the elastic and the inertial force, a fluidic inductance  $L_{fl}$  and a fluidic capacity  $C_{fl}$  need to be defined in a suitable way. If, e.g., a membrane is

**Fig. 193** Capillary with a membrane (a) and analog electrical circuit (b)



**Table 20** Devices of different physical domains which are analog to each other

Fluidics	Electronics	Mechanics
Pump	Voltage source	Engine
Passive valve	Diode	Ratchet
Active valve	Transistor	Brake
Hydraulic press	Transformer	Gear or lever

build in a flow channel as shown in Fig. 193a, the interrelationship between deflection and pressure difference is described by (53) (page 38). If the deflections  $w_0$  are small compared with the thickness  $d_M$  of the membrane, the third term in the parenthesis can be neglected. The deflection of the membrane can be expressed by the displaced volume  $V$  employing (303) (page 170) resulting in:

$$\Delta p = \frac{8 d_M}{\pi R_M^4} \left( \frac{4}{3} \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 \right) V. \tag{425}$$

The analogy in (422) results in:

$$v = \frac{\partial x}{\partial t} \leftrightarrow \Phi_F = \frac{\partial V}{\partial t} \Rightarrow x \leftrightarrow V. \tag{426}$$

Comparing the above equation with the equations of the elastic force in Table 19 now yields the definition of the fluidic capacity  $C_{fl}$ :

$$F_k = kx \leftrightarrow U = \frac{Q_{el}}{C_{el}} \leftrightarrow \Delta p = \frac{V}{C_{fl}}. \tag{427}$$

Comparing (425) and (427) now yields the fluidic capacity  $C_{fl}$  of the given problem:

$$C_{fl} \approx \frac{\pi R_M^4}{8 d_M} \frac{1}{(4/3) (d_M^2/R_M^2) (E_M/1 - \nu_M^2) + \sigma_0}. \tag{428}$$

If different physical domains are to be coupled, the energy flow leaving one domain needs to be the same as the energy flow entering the other domain. A simple

example is the electroosmotic actuator (cf. page 122f). A volume flow  $\Phi_F$  is generated by a voltage  $U$  applied to a capillary filled with a polar liquid such as water. The energy flow, i.e., the power, of the electric  $P_{el}$  and the fluidic  $P_{fl}$  domain, respectively, are:

$$P_{el} = U I_{el} = P_{fl} = \Delta p \Phi_F. \tag{429}$$

The electroosmotic actuator can be drawn as an equivalent circuit together with the electrical circuit of the electronics as shown in Fig. 194a. The electrical current  $I_{el}$  and the flow  $\Phi_F$  can be expressed by voltage  $U$  and electrical resistance  $R_{el}$  and pressure difference  $\Delta p$  and fluidic resistance  $R_{fl}$ , respectively:

$$P_{el} = \frac{U^2}{R_{el}} \text{ and } P_{fl} = \frac{\Delta p^2}{R_{fl}} \Rightarrow R_{el} = \frac{U^2}{\Delta p^2} R_{fl}. \tag{430}$$

The pressure difference  $\Delta p_{\text{EOF}}$  driving the fluid as a function of the voltage is calculated with (214) (page 124):

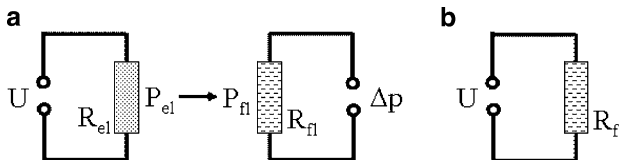
$$\Delta p_{\text{EOF}} = \frac{32}{D_h^2} \epsilon_0 \epsilon_r U \zeta. \tag{431}$$

Inserting the above equation and the flow resistance  $R_{fl}$  [(178) on page 108] into (430), yields the electrical resistance of the electroosmotic actuator:

$$R_{el} = \frac{D_h^2 \eta L_E}{32 A \epsilon_0^2 \epsilon_r^2 \zeta^2}. \tag{432}$$

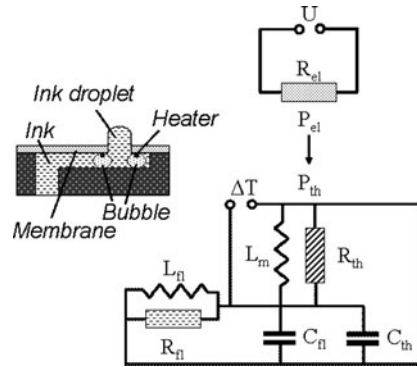
The electroosmotic actuator may also be drawn as a simplified schematic equivalent circuit which shows the electric and fluidic circuit combined and sometimes allows an easier understanding of the interrelationships (cf. Fig. 194b).

In general, the equivalent circuits of microsystems are much more complex than in the case of the electroosmotic actuator. The head of an ink-jet printer shown in Fig. 189 (page 252) could be depicted by an equivalent circuit (cf. Fig. 195) showing the electrical resistance of the heater  $R_h$ , voltage  $U$  applied to the heater, temperature change generated by the heater  $\Delta T$ , accelerated mass  $m_F \leftrightarrow L_m$  of the



**Fig. 194** Coupled electrical and fluidic circuit (a) and simplified combined circuit (b)

**Fig. 195** Cross-section and equivalent circuit of an ink-jet printer



droplet ejected from the printer, thermal resistance  $R_{th}$  describing the heat loss to the environment, fluidic capacity  $C_{fi}$  of the bubble, thermal capacity  $C_{th}$  of the bubble and the liquid ink, flow resistance  $R_{fi}$  of the channel to the reservoir, and fluidic inductance  $L_{fi}$  of the accelerated liquid ink in the channel.

Such an equivalent circuit may help to understand the interrelationships of the components of the microsystem. For example, it becomes clear that the mass and flow resistance of the liquid in the channel to the reservoir should be designed much larger than the mass of a droplet. From the response time of the circuit, it can be estimated how large the thermal resistance and fluidic and thermal capacitance should be to achieve ejection of a droplet of the desired size in the desired time. However, it remains to be difficult to calculate the time response of a complex circuit. Fortunately, there are computer programs available which have been developed to simulate electronic circuits. Such programs can be employed to calculate, e. g., the time response of the equivalent circuit of the ink-jet printer, and, that way, solve the problem of the microsystems comprising several physical domains.

Besides the above, analogies can help to find a solution for a problem. If there is a problem in one physical domain; e.g., a force is too weak, it may help to have a look on the same problem in another physical domain. A force in the mechanical domain can correspond to a voltage in the electronic domain. If a voltage is too small, it can be enhanced with a transformer which results in less current at more voltage. Now the analogy may go back into the mechanical domain and it may be found that a lever in the mechanical domain corresponds to a transformer in the electrical domain. In the simple case described here, certainly it was clear that a lever could solve the problem but changing the point of view by moving to another analog domain may help also to solve more complex problems.

The analogy of devices from different physical domains is not limited to the case of a transformer and a lever. Table 20 shows a list of analog devices which, e.g., have analog characteristic curves.

Table 21 shows possible analogies.



**Table 21** Analog quantities in different physical domains. Other analogies between the same quantities are also possible as a function of the given problem

	Electronics	Mechanics	Rotation	Fluidics	Thermodynamics
Potential	U	F	$M_t$	$\Delta p$	T
Status	$Q_{el}$	x	$\varphi$	V	S
Flow	$I_{el}$	v	$\omega$	$\Phi_F$	$\dot{S}$
	$\dot{I}_{el}$	$a_m$	$\dot{\omega}$	$\dot{\Phi}_F$	
	$R_{el}$	$D_R$	$D_R$	$R_{fl}$	$R_{th}$
	$C_{el}$	1/k	1/k	$C_{fl}$	$C_{th}$
Energy flow	$L_{el}$	$m_K$	$I_m$	$L_{fl}$	
	$U I_{el}$	F v	$M_t \omega$	$\Delta p \Phi_F$	$T \dot{S} = \dot{Q}_{th}$
Energy	$U Q_{el}$	F x	$M_t \varphi$	$\Delta p V$	$T S = Q_{th}$

The denominations are found on page viif

## Exercises

### Problem 43

A microfluidic filter as shown in Fig. 191 (page 255) consists of several simple capillaries as shown in Fig. E32. The dimensions of the capillaries are shown in the figure. All channels are 1 mm in height. When water (viscosity: 1 mPa s) is entering into this capillary system at a pressure difference of 3 Pa, what is the volume flow at the outlet? Please perform the calculation as if the flow would go straight through the channels. In reality, at each corner, there is an additional flow resistance which does not enter into the calculation.

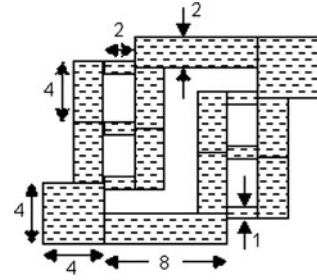
The so-called delta-wye conversion in electrotechnique is shown in Fig. E33. It can help to solve the problem, but is not necessarily needed.

### Problem 44

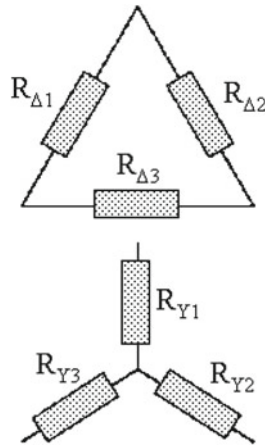
A micropump does not produce a continuous pressure but superimposes it with a sine wave pressure fluctuation. This is inappropriate for the intended application. Therefore, the pressure generated shall be made smoother by designing downstream of the pump a constriction and a membrane. Figure E34 shows the constriction and the membrane together with an electronic equivalent circuit and the equation describing the voltage change due to the circuit.

The constriction is 3 mm, 100  $\mu\text{m}$ , and 100  $\mu\text{m}$  in length, height, and width, respectively. The fluid is water, i.e., the viscosity is 1 mPa s. Radius, thickness, Poisson’s ratio, and Young’s modulus of the membrane are 1 mm, 10  $\mu\text{m}$ , 0.5, and 2 MPa, respectively, and there is no residual stress.

**Fig. E32** A system of microchannels. All measures are given in millimeters



**Fig. E33** Delta-wye configuration in electrotechnique

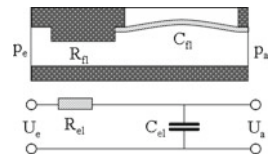


$$R_{Y1} = \frac{R_{\Delta 1} R_{\Delta 2}}{R_{\Delta 1} + R_{\Delta 2} + R_{\Delta 3}}$$

$$R_{Y2} = \frac{R_{\Delta 2} R_{\Delta 3}}{R_{\Delta 1} + R_{\Delta 2} + R_{\Delta 3}}$$

$$R_{Y3} = \frac{R_{\Delta 3} R_{\Delta 1}}{R_{\Delta 1} + R_{\Delta 2} + R_{\Delta 3}}$$

**Fig. E34** Fluidic system and electronic equivalent circuit



$$\frac{U_a}{U_e} = \frac{1}{\sqrt{R_{el}^2 \omega^2 C_{el}^2 + 1}}$$

- (a) Calculate the pressure fluctuations expected behind the membrane when their frequency and amplitude are 1 Hz and 200 Pa, respectively, before the constriction.
- (b) As a consequence of the constriction, the flow rate is much reduced. Experiments show that a width and a height of the constriction should be 500  $\mu\text{m}$ . Please calculate how large the circular membrane needs to be to achieve the same smoothening effect as in (a).

# Mechanical Devices for Electronics

*Electromechanical switches*, also called relays, are devices which are designed to switch an electrical current by a voltage employing some mechanical means. For most applications, currents are switched by a transistor and this is the cheapest known way to control the flow of a current.

However, for certain applications, transistors are not suitable. For example, high frequency signals cannot be switched off by a transistor. Transistors are a kind of capacitor from the point of view of the signal which can overcome the distance between source and drain, if its frequency is high enough. Other examples are devices which for safety reasons are not allowed to be realized as a microelectronic circuit. For example, the switch stopping a robot when a person enters into its working range in some countries must be a mechanical switch.

In the past, relays have been employed to provide a safe interruption of an electrical circuit. Microtechnique allows constructing *relays* which are smaller, lighter, and cheaper. Figure 196 shows how an electromechanical switch may look as [70]. A beam is attracted by the electrostatic force of the control signal and closes an electrical circuit.

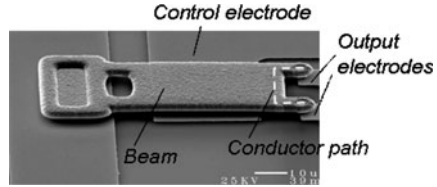
If the output line would be realized just by one contact of the beam and a counter electrode on the substrate, the beam could be deflected down by a large voltage at the output. Therefore, a small bridge is closing the contacts between the output electrodes when the beam is deflected down by the voltage applied to the control electrode. This design also reduces capacitive coupling between the output electrodes.

If the distance of the control electrode with area  $A_C$  from the non-deflected beam is  $d_0$ , the electrostatic force  $F_C$  pulling down a beam as a function of the deflection  $w_0$  is according to (222) (page 131):

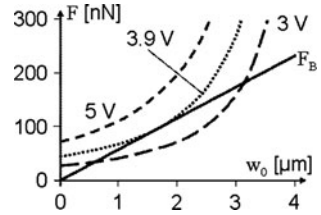
$$F_C = \frac{1}{2} \frac{\epsilon_0 \epsilon_r A_C}{(d_0 - w_0)^2} U^2. \quad (433)$$

The elastic force  $F_B$  of a beam clamped at one end is described by (97) on page 66. Electrostatic and elastic forces are shown in one graph in Fig. 197 as a function of the deflection  $w_0$  of a beam with a Young's modulus  $E_B$ , distance  $L_B$

**Fig. 196** Electromechanical switch [70]. © [1997] IEEE



**Fig. 197** Elastic force  $F_B$  of a beam and electrostatic force  $F_C$  pulling it down as a function of deflection  $w_0$  and applied voltage



between the center of the control electrode and the fixation point of the beam, width  $b_B$ , and thickness  $d_B$  of 120 GPa, 550  $\mu\text{m}$ , 40  $\mu\text{m}$ , and 2  $\mu\text{m}$ , respectively, which is mounted at  $d_0 = 5 \mu\text{m}$  above the control electrode.

Figure 197 shows that at a low voltage such as 3 V the electrostatic force is larger than the elastic force of the beam only up to a deflection of 0.6  $\mu\text{m}$ . As a consequence, there is a stable equilibrium of forces at that deflection and the beam would not move beyond that point. There is another equilibrium of forces at 3.1  $\mu\text{m}$  which is unstable. That is, a beam deflected more than 3.1  $\mu\text{m}$  will bend completely down and a beam deflected less will go back to the stable position at 0.6  $\mu\text{m}$ .

When the voltage is larger than 3.9 V, there is no stable position for the beam and it is completely bent down. The voltage necessary to bend the beam completely down is called the *pull-in-voltage*. Obviously, this is the desired situation and an algorithm is required to find the pull-in-voltage as a function of the geometrical and material parameters of the beam.

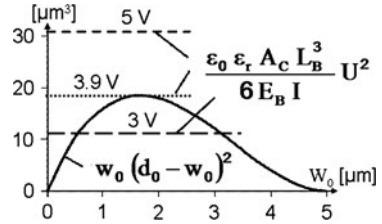
At the equilibrium of forces, the electrostatic force according to (433) and the elastic force of the beam [(97) on page 66] are equal:

$$F_C = \frac{1}{2} \frac{\epsilon_0 \epsilon_r A_C}{(d_0 - w_0)^2} U^2 = F_B = 3 \frac{E_B I}{L_B^3} w_0$$

$$\Rightarrow w_0(d_0 - w_0)^2 = \frac{\epsilon_0 \epsilon_r A_C L_B^3}{6 E_B I} U^2. \tag{434}$$

Both sides of the above equation are drawn in Fig. 198 for the parameters used in Fig. 197. The left term is a polynomial and the right one is a constant with respect to the deflection. The crossing points of the two curves mark the equilibrium of

**Fig. 198** Left and right side of (434)



forces. The pull-in-voltage is achieved when the constant described by the right side of (434) is larger than the maximum of the left side of the equation.

The maximum of the left side of (434) is found by differentiating it with respect to the deflection, calculating the zero, and inserting it into (434) again. This way, the pull-in-voltage  $U_p$  is found:

$$\frac{4}{27} d_0^3 = \frac{\epsilon_0 \epsilon_r A_C L_B^3}{6 E_B I} U_p^2 \Rightarrow U_p = \frac{2}{3} d_0 \sqrt{2 \frac{E_B I d_0}{\epsilon_0 \epsilon_r A_C L_B^3}}. \tag{435}$$

When the switch is closed, a voltage  $U_{min}$  smaller than  $U_p$  is sufficient to keep it closed.  $U_{min}$  can be calculated from the equilibrium of elastic and electrostatic forces of the closed switch. The distance  $d_i$  of the control electrodes of a closed switch need to be larger than zero because a short-circuit must be avoided. The corresponding equation of (434) for a closed switch is:

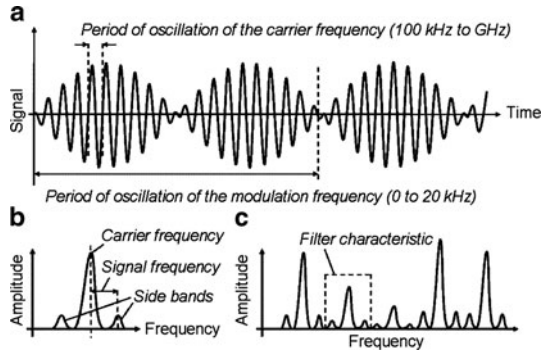
$$F_C = \frac{1}{2} \frac{\epsilon_0 \epsilon_r A_C}{d_i^2} U_{min}^2 = F_B = 3 \frac{E_B I}{L_B^3} (d_0 - d_i) \\ \Rightarrow U_{min} = d_i \sqrt{6 \frac{E_B I}{\epsilon_0 \epsilon_r A_C L_B^3} (d_0 - d_i)}. \tag{436}$$

If the voltage becomes smaller than  $U_{min}$ , the switch will not open completely but go to its stable position. To completely open the switch, the voltage at the control electrode needs to be set back to zero.

*Electromechanical filters* are another kind of mechanical devices employed in electronics. Devices for mobile communication such as cellular phones, radios, television, WLAN, and communication satellites need to filter out a certain frequency band on which the desired information is sent.

There are several ways how to modulate a carrier frequency to transport information with it. Figure 199a shows an electromagnetic signal sent at a carrier frequency and modulated with a signal frequency which contains the information, e.g., an acoustic signal. The envelope curve of the curve shown in Fig. 199a describes the signal. A Fourier transformation of the signal shown in Fig. 199a is displayed in Fig. 199b. The carrier frequency is surrounded by the so-called side

**Fig. 199** (a) Transmitted signal, (b) frequency spectrum of the signal, and (c) frequency spectrum as received by an antenna and necessary filter characteristic to separate the desired information



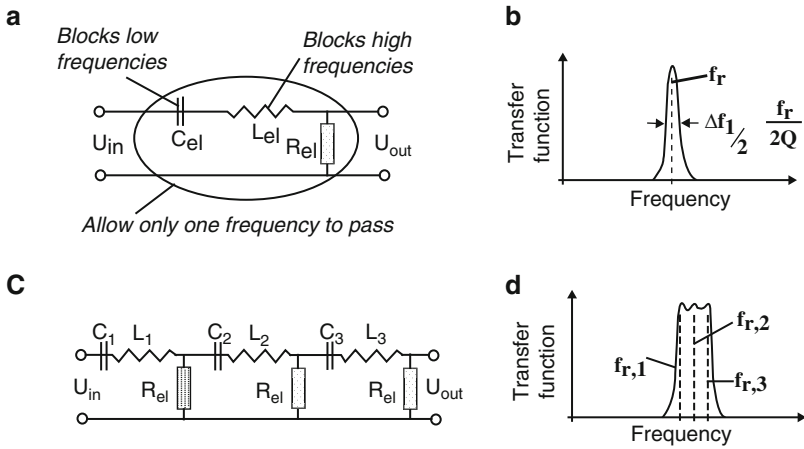
bands which contain the desired information. Different information are broadcast simultaneously by a lot of stations at different carrier frequencies resulting in a lot of frequencies at which information is transmitted (Fig. 199c). There are techniques available which allow transmitting only one side band suppressing the carrier frequency and the second side band. This allows to send simultaneously more information in a given frequency range. Transmitting digital information results in much narrower frequency bands because only the on and off state of the carrier frequency needs to be detected.

In any case, the receiver of a communication device needs to be tuned to a certain frequency band to record only the desired information. Therefore, some band pass filter is required which filters out the desired frequency band. The sharper the band pass filter is the more information can be transmitted simultaneously at neighboring frequencies.

An electronic filter is shown schematically in Fig. 200a. The capacitor suppresses low frequencies, while the coil suppresses high frequencies. As a result, there is a certain frequency called the resonance frequency  $f_r$  which is passing the filter best while other frequencies are suppressed. Figure 200b shows the transfer function of the filter shown in Fig. 200a, which describes the ratio of the voltage amplitudes of a sine signal at the output  $U_{out}$  and at the input  $U_{in}$  as a function of their frequency.

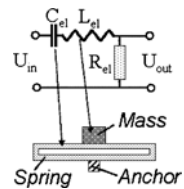
The transfer function in Fig. 200b does not show the desired form, because it filters out only one resonance frequency. Instead, it is necessary to provide a circuit with a rectangular transfer function, which allows all frequencies in a certain range to pass the filter and suppresses all other frequencies. Such a transfer function is approximated by a serial connection of several filters with neighboring resonance frequencies. Figure 200c shows the schematic drawing of a filter composed of three individual filters, and Fig. 200d shows their transfer function which contains the three individual resonances frequencies.

The resonance frequency of an electronic filter can be calculated with (420) (page 254), which also describes the resonance frequency of a system composed of a mass and a spring. That is, instead of an electronic filter containing a capacitor and



**Fig. 200** (a) Electronic filter, (b) transfer function of the filter shown in (a), (c) electronic band pass filter, and (d) transfer function of the band pass filter

**Fig. 201** Electronic and mechanical resonator



a coil, a mechanical vibrating structure can be employed which comprises a spring and a mass as schematically shown in Fig. 201. The resistor in the electronic filter corresponds to the damping of the moving mass of the mechanical system.

A difference between the electronic and the mechanical filter is that the mechanical filter is only a weak function of temperature (because Young’s modulus is a function of temperature) and damping can be reduced more by suitable designs and evacuating the filter housing. Therefore, with mechanical filters, a transfer function with sharper edges can be constructed. Besides this, the resonance frequency of a mechanical filter can be adjusted by simply applying a voltage (see below).

A measure of the sharpness of the edges of the transfer function of a filter is the so-called quality factor  $Q$  which is defined as  $2\pi$  times the ratio of the potential energy per period of the oscillator and the energy loss per period. The quality factor is related to the angular frequency at the resonance  $\omega_r$  and the damping constant  $D_R$ :

$$Q = \frac{\omega_r}{D_R} \approx \frac{\sqrt{k m_K}}{D_R} \text{ with } F_R = D_R \text{ v.} \tag{437}$$

Electronic filters have quality factors of some hundred, while micromechanical filters achieve more than 10,000.

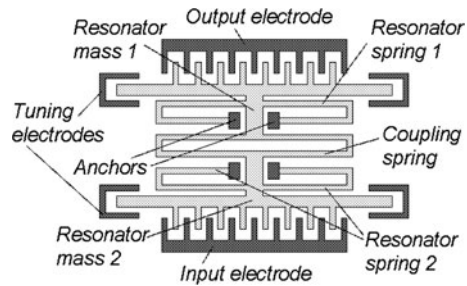
To excite a mechanical resonator composed of a spring and a mass by an electrical signal, an electrostatic actuator is a simple solution. Such an arrangement is shown in Fig. 202. It consists of two mechanical resonators coupled by an intermediate spring.

An alternating voltage applied to the input electrode results in a mechanical vibration of the resonator system. It is necessary to apply a bias voltage to the resonator because otherwise the resonator would be attracted both by the positive and the negative half-wave of the signal resulting in a vibration with the doubled frequency. The bias voltage is also required to generate a signal at the output electrode when the capacity of the output is changing due to the moving resonator.

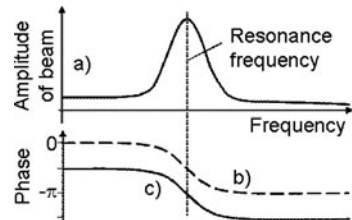
As shown in Fig. 203a, the resonator is excited to larger vibrations at its resonance frequency, and therefore out of a frequency mix at the inlet mainly the resonance frequency is transmitted to the output. A combination of two (cf. Fig. 202) or more resonators at adjacent resonance frequencies results in a transfer function which is close to the desired one. At frequencies far below the resonance of the mechanical oscillator, there is no phase shift between the input signal and the vibration, and as a consequence there is also no phase shift between the input and the output voltage. However, at the resonance, there is a phase shift of  $-\pi/2$  and at higher frequencies it is  $-\pi$  (cf. Fig. 203b).

At the output, there is a capacitance  $C_{el}$  varying with the frequency and phase of the resonator. The constant bias voltage  $U_B$  is applied to the capacitor. In general, the electrical current  $I_{el}$  generated at a capacitor is calculated from the charge change as a function of time:

$$Q_{el} = C_{el}U \Rightarrow \frac{\partial Q_{el}}{\partial t} = I_{el} = \frac{\partial C_{el}}{\partial t} U_B. \tag{438}$$



**Fig. 202** Electromechanical filter consisting of two resonators. All *bright structures* are cantilevers



**Fig. 203** Schematic drawing of amplitude and phase of beam vibrations



The capacitance of comb electrodes is a linear function of their overlapping length [cf. (223) on page 132], and therefore varies with the same phase and relative amplitude as the deflection of the resonator. Denoting the capacity of the output electrode in the idle state of the resonator by  $C_0$  and the maximum capacity change due to the vibrating beam by  $\Delta C$  it is obtained:

$$\begin{aligned} C_{el} &= C_0 + \Delta C \sin(\omega t + \varphi) \Rightarrow I_{el} = \frac{\partial C_{el}}{\partial t} U_B \\ &= \Delta C \omega \sin\left(\omega t + \varphi - \frac{\pi}{2}\right) U_B. \end{aligned} \quad (439)$$

In the above calculation, it has been taken into account that the derivative of a sine function is the same as a shift by  $-\pi/2$ . Thus, the current leaving the output electrode has a phase shift of  $-\pi/2$  plus the phase  $\varphi$  of the beam. Figure 203c shows the phase of the output current. The current signal then can be converted into a voltage, in the easiest case by a resistance mounted in series to the output electrode.

The amplitude of the current signal is proportional to  $(\Delta C \omega)$  and  $\Delta C$  is a function of  $\omega$  because the amplitude of the beam vibration is a function of the frequency  $\omega/(2\pi)$  as shown in Fig. 203a. As a consequence, the maximum of the transfer function is at a frequency which is a bit larger than the resonance frequency of the beam.

The resonance frequencies of the resonators can be tuned by additional electrodes such as shown in Fig. 202. When the resonators are not deflected, they are in an unstable equilibrium (cf. Fig. 96 on page 132). The force acting on the resonator is composed of the elastic force of the beams which can be expressed by their spring constant  $k_m$  and the electrostatic force described by (225) (page 133):

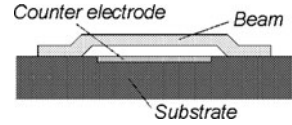
$$F_B = -k_m x + \frac{1}{2} \varepsilon_0 \varepsilon_r A_C U^2 \left( \frac{1}{(d_C - x)^2} - \frac{1}{(d_C + x)^2} \right). \quad (440)$$

The deflections  $x$  of electromechanical filters usually are small. Therefore, it is a good approximation to calculate the Taylor expansion at  $x = 0$ :

$$\begin{aligned} F_B &= F_B(x) + \left. \frac{\partial F_B}{\partial x} \right|_{x=0} x = -k x + 2\varepsilon_0 \varepsilon_r A_C U^2 d_C x \\ &= (-k_m + k_{el})x \quad \text{with } k_{el} := 2\varepsilon_0 \varepsilon_r A_C U^2 d_C. \end{aligned} \quad (441)$$

In the above equation,  $k_{el}$  is the so-called electrical spring constant. By applying a voltage to the tuning electrodes, the spring constant of the resonators can be diminished and thus the resonance frequency of the resonators [calculated with (420) on page 254] and the transfer function of the filter can be tuned in the desired way. The tunable range is limited because too large deflections of the beam result in a nonlinear transfer function and the electrodes should not touch each other and

**Fig. 204** Cross-section of a vibrating beam employed as an electromechanical filter



generate a short-circuit. It needs to be noted that the bias voltage applied to inlet and outlet electrodes also results in an electrical spring constant which has to be taken into account when the resonance frequency is calculated.

A simpler design which can be employed as an electromechanical filter is just a beam mounted above an electrode (cf. Fig. 204). As described above, it is necessary to apply some bias voltage  $U_B$  to the beam to achieve a beam vibration with the same frequency as the input voltage. The beam together with the counter electrode forms a capacitor with a capacity which is a function of frequency. In this case, the calculation of the transfer function is more complex than for the filter shown in Fig. 202 because both the capacity  $C_{el}$  and the voltage  $U$  are a function of time:

$$Q_{el} = C_{el}U \Rightarrow \frac{\partial Q_{el}}{\partial t} = I_{el} = \frac{\partial C_{el}}{\partial t}U + C_{el} \frac{\partial U}{\partial t}. \quad (442)$$

The capacity consists of two parts: The constant capacity of the beam in its idle state  $C_0$  and a part varying as a function of time with the frequency  $f = \omega/(2\pi)$  and phase  $\varphi$  of the beam vibrations. The capacity is not a linear function of the deflection of the beam because it is moving normal to the electrode. However, if the deflection is small compared with the beam's distance to the counter electrode in the idle state, the varying part of the capacity can be approximated by a sine function with an amplitude  $\Delta C$  which is a function of the amplitude of the beam vibrations:

$$C_{el} = C_0 + \Delta C(\omega) \sin(\omega t + \varphi). \quad (443)$$

The voltage is the sum of the constant bias  $U_B$  and the varying part of the input signal. Beam vibrations are generated by the input signal, therefore the frequency of beam vibration and input are the same:

$$U = U_B + \Delta U \sin(\omega t). \quad (444)$$

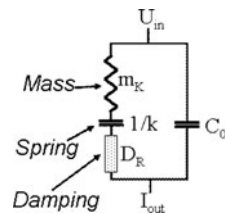
Inserting the above (443) and (444) into (442) yields (Again the derivative of a sine function is a phase shift of  $-\pi/2$ ):

$$I_{el} = \omega \Delta C \sin\left(\omega t + \varphi - \frac{\pi}{2}\right) U_B + \omega \Delta C \sin\left(\omega t + \varphi - \frac{\pi}{2}\right) \Delta U \sin(\omega t) + C_0 \omega \Delta U \sin\left(\omega t - \frac{\pi}{2}\right) + \Delta C \sin(\omega t + \varphi) \omega \Delta U \sin\left(\omega t - \frac{\pi}{2}\right). \quad (445)$$

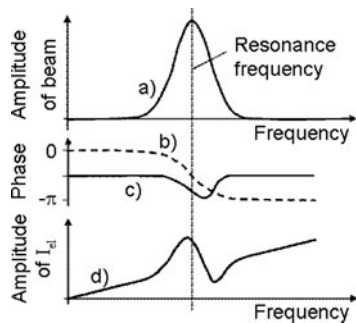
In the above equation, the terms including the product of  $\Delta C$  and  $\Delta U$  can be neglected when they are small compared with  $C_0$  and  $U_B$ , respectively. The second of the remaining two terms describes the current generated at a constant capacity. The first term describes the effect of the vibrating beam. Therefore, the equivalent circuit of the device approximately can be drawn as the capacity  $C_0$  of the beam in its idle state and parallel to this the vibrating beam described by an inductance ( $L_m = m_K$ ), a capacitor ( $C_m = 1/k_m$ ), and a resistor ( $R_m = D_R$ ) (cf. Fig. 205).

However, this is only an approximate view and the resonance frequency should be calculated by a numerical analysis of (445). Figure 206 shows the result of an analytical calculation of the amplitude and phase of  $I_{e1}$  with (445) as a function of frequency. For  $\Delta C$ , the deflection of a vibrating beam was approximated by a Gaussian curve around the resonance frequency  $f_r$  and a width of  $2 f_r$  as shown in Fig. 206a. (It is the advantage of electromechanical filters that the width is much less than shown here. However, the relative position of the maxima can be recognized more clearly at a larger width.) The phase of the vibrating beam was approximated by a hyperbolic tangent function as shown in Fig. 206b. The shape of the curve describing the amplitude of the output current (Fig. 206d) is a function of the ratio of the products of  $[U_B \Delta C(\omega)]$  and  $[C_0 \Delta U]$ . For the calculation shown in Fig. 206, the ratio  $U_B/C_0$  was 1.7. Far from the resonance,  $\Delta C(\omega)$  is small and the current output is only the one of a capacitor which is increasing proportional to frequency.

**Fig. 205** Equivalent circuit of a vibrating beam employed as an electromechanical filter



**Fig. 206** Assumed amplitude (a) and phase (b) of the vibrating beam shown in Fig. 204; and amplitude (d) and phase (c) of the current output calculated with (445)



In the near of the resonance, a maximum and a minimum occur below and above the resonance frequency of the beam. The phase of the output signal (Fig. 206c) is shifting to  $-\pi$  at the beam resonance, while at other frequencies it is  $-\pi/2$  as expected for a capacitor. The output current shown in Fig. 206d is not suitable as a band pass filter. However, a combination of such devices can result in a transfer function with very steep slopes. The overall rise of the signal due to the constant capacity  $C_0$  can be eliminated with an additional electronic filter which does not require a sharp and temperature stable edge. It needs to be noted that there is an electrical spring constant also for a simple beam over an electrode, and therefore, the resonance will be changed by and can be tuned with the bias voltage.

It is the purpose of electromechanical filters to obtain sharp edges of the transfer function. Therefore, damping of the mechanical vibration should be reduced as much as possible. Damping is mostly due to friction of the oscillating structure in the surrounding air and also due to sound emission. As a consequence, it is usual to mount the device in an evacuated housing.

## Exercises

### *Problem 45*

Please calculate the minimum switching voltage for the electrostatic switch shown in Fig. 196 (page 262). Do not take into account, the orifice in the beam in the near of its fixation. The dimensions and properties of the beam required for the calculation are shown in the following table:

Young's modulus of beam	120 GPa
Width of beam	20 $\mu\text{m}$
Thickness of beam	2 $\mu\text{m}$
Length of switching electrode	40 $\mu\text{m}$
Width of switching electrode	15 $\mu\text{m}$
Distance between beam fixation and end of switching electrode	50 $\mu\text{m}$
Relative permittivity	1
Distance between beam in idle state and substrate	1 $\mu\text{m}$
Thickness of isolation layer between touching electrodes	500 nm

- Calculate also how much the voltage needs to be lowered again to open the relay and at what acceleration the switch will open and close, respectively.
- The result of (a) indicates that the orifice near the fixation of the beam is absolutely necessary. What could also be done to enhance the movability of the beam?
- What is the electrical capacity of the signal line of the open relay when the overlapping area is  $2 \times 2 \mu\text{m}^2$ ?

**Problem 46**

The resonance frequency of an electromechanical filter can be tuned by the bias voltage.

- (a) Calculate the resonance frequency of such a filter made of a beam fixed at both ends. Perform the calculation both with and without bias voltage. The required quantities are given in the following table:

Relative permittivity	1
Absolute permittivity	$8.9 \times 10^{-12}$ A s/(V m)
Young’s modulus of beam	120 GPa
Residual stress of beam	0 MPa
Width of beam	100 $\mu\text{m}$
Length of beam	200 $\mu\text{m}$
Thickness of beam	10 $\mu\text{m}$
Density of the spring	$2.33 \times 10^3$ kg/m <sup>3</sup>
Distance between beam in idle state and substrate when no voltage is applied	1 $\mu\text{m}$
Distance between beam in idle state and substrate when bias voltage is applied	0.77 $\mu\text{m}$
Inner area of one capacitor plate	$100 \times 100$ $\mu\text{m}^2$
Bias voltage	135 V

- (b) Please calculate the quality factor of the filter in (a) due to the squeeze-film effect of the air in the gap between beam and substrate. Assume a spring constant of 10,000 Pa m, a beam mass of 0.4  $\mu\text{g}$ , and a dynamic viscosity of air of 18.2  $\mu\text{Pa s}$ . As an approximation, assume that the entire beam is moving up and down without bending at a distance of 0.77  $\mu\text{m}$  to the substrate.

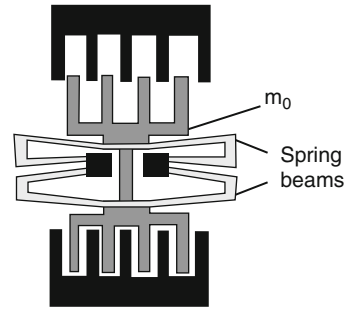
**Problem 47**

Please calculate the quality factor and the resonance frequency of the filter shown in Fig. E35. Neglect the mass of the spring beam for the calculation of the quality factor but not for the calculation of the resonance frequency.

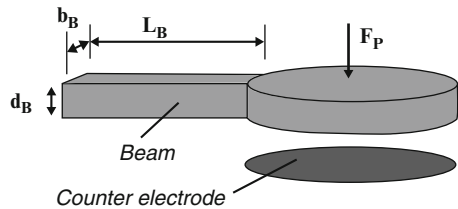
The properties required for the calculation are found in the following table:

Young’s modulus of beam	120 GPa
Width of beam	10 $\mu\text{m}$
Length of beam	200 $\mu\text{m}$
Thickness of beam	10 $\mu\text{m}$
Density of the material	$2.33 \times 10^3$ kg/m <sup>3</sup>
Distance to the substrate	2 $\mu\text{m}$
Viscosity of air	18.2 $\mu\text{Pa s}$
Area of movable structures	$2.50 \times 10^{-2}$ mm <sup>2</sup>
Mass of beams $m_B$	163 ng
Central mass $m_0$	419 ng

**Fig. E35** Electromechanical filter. Only the *black areas* are fixed to the substrate



**Fig. E36** Acoustical microswitch



**Problem 48**

An acoustical microswitch shall close an electrical contact when a certain frequency appears. For this, a beam is fabricated with a circular area, 2 mm in diameter, at its end (cf. Fig. E36). A voltage can be applied between the beam and a counter electrode. The beam shall be deflected by an acoustic signal so much that it touches the counter electrode and, this way, generates a short-circuit.

The quantities required for the calculation are given in the following table:

Relative permittivity	1
Absolute permittivity	$8.9 \times 10^{-12}$ A s/(V m)
Young's modulus of beam	120 GPa
Residual stress of the beam	0 MPa
Width of beam	300 $\mu$ m
Length of beam	3 mm
Thickness of beam	10 $\mu$ m
Density of beam	$2.33 \times 10^3$ kg/m <sup>3</sup>
Distance between beam and counter electrode when no voltage is applied	2 $\mu$ m
Voltage applied to the beam	0.1 V

- (a) A microphone measures the pressure of an acoustical wave emitted by a source. At a distance of 0.5 m from the source, the pressure is 1 Pa. It reduces proportional to the inverse of the distance. When the frequency of the acoustic wave is not the same as the resonance frequency of the beam, as an approximation it can be

assumed that the acoustical pressure is applied once to the area at the end of the beam.

What force  $F_p$  is acting at the end of the beam when it is placed 7 m away from the source? Is this force alone (with no voltage applied) enough to deflect the beam so much that it touches the counter electrode?

- (b) A voltage applied between the end of the beam and the counter electrode helps to close the switch. Is the switch closed with the help of the electrostatic force? Verify that the switch is closed by the additional electrostatic force just 70 nm.
- (c) When the beam is excited at its resonance frequency, its amplitude is significantly enhanced and the switch is closed. What resonance frequency has to be expected? (Ignore squeeze-film effect and electrostatic force.)
- (d) Why is only such a small voltage applied to the beam?

# Pressure Sensors

Pressure sensors play an important role in microsystem industries. They are employed in a lot of different application such as triggering for side air bags and monitoring tire pressure in cars, detecting the elevation for hiking or paragliding, and flow measuring (cf. page 299f).

The main advantages of miniaturized pressure sensors are low cost, small size and weight, and quick response time in pneumatic applications. A pressure change cannot be measured before the fluid has adopted the change. This is delayed in compressible fluids because some fluid needs to flow into or off the sensor and this takes some time. The time required to change the pressure in the dead volume  $V_T$  of a pressure sensor as a function of the flow resistance  $R_{fl}$  of the feed channel and the final pressure  $p_e$  can be approximated with (381) on page 218. Obviously, the small dead volume of a miniaturized pressure sensor is an advantage and the flow resistance of the feed channel needs to be designed as large as possible.

Pressure sensors contain a membrane deflected by the pressure difference to be measured. The deflection of a membrane as a function of the pressure difference is calculated with the equations shown in Table 2 (page 50). Equation (53) (on page 38) describes the interrelationship between pressure and deflection of a circular membrane in general:

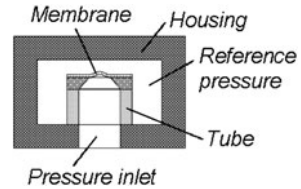
$$\Delta p = \frac{4 d_M w_0}{R_M^2} \left( \frac{4}{3} \frac{d_M^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} + \sigma_0 + \frac{64}{105} \frac{w_0^2}{R_M^2} \frac{E_M}{1 - \nu_M^2} \right). \quad (53)$$

For noncircular membranes, the interrelationship between pressure difference  $\Delta p$  and deflection  $w_0$  is described by a similar equation [cf. (67), page 46]. Therefore, in the following, a circular membrane is assumed representative for all membrane shapes.

The first, second, and third term in the parenthesis of (53) describe the effects of bending moments, residual stress, and stress due to straining, respectively. The second term describing the effect of the residual stress is disturbing because the stress of a micromembrane is easily changed by outer forces acting on the sensor housing. Therefore, it is usual to mount the membrane at the free end of a tube where the stress of the housing is kept away from the sensing membrane (cf. Fig. 207).



**Fig. 207** Schematic cross-section of a pressure sensor mounted at the free end of a tube



The third term in the parenthesis of (53) represents the stress change of the neutral fiber (cf. page 5) and the entire membrane due to the straining of the membrane generated by its deflection. This term is responsible for a nonlinear interrelationship between pressure difference and membrane deflection. Therefore, for a linear characteristic curve, it is required that the deflection  $w_0$  of the membrane is much less than its thickness  $d_M$ . As a consequence, the third term in the parenthesis is much smaller than the first one, and, the deflection is a linear function of the pressure difference.

The deflection of the membrane can be detected by strain gauges, as a capacity change, or the frequency change of a resonating microstructure. When strain gauges are employed, it is necessary to compensate for changes of their electrical resistance as a function of temperature. This is accomplished by employing a bridge (cf. Fig. 39 on page 54).

The potentiometer is adjusted such that the output voltage  $U_m$  of the bridge is zero when the membrane is not deflected. Therefore, the output voltage is equal to the change of the voltage drop over one of the resistance, e.g.  $R_2$ :

$$\begin{aligned} U_2 = \frac{R_2}{R_1 + R_2} U_0 &\Rightarrow U_m = \frac{\partial U_2}{\partial R_1} \Delta R_1 + \frac{\partial U_2}{\partial R_2} \Delta R_2 \\ &= U_0 \left( \frac{-R_2 \Delta R_1}{(R_1 + R_2)^2} + \frac{R_1 \Delta R_2}{(R_1 + R_2)^2} \right). \end{aligned} \quad (446)$$

If  $R_1$  equals  $R_2 = R_{el}$ , the above equation results in:

$$U_m = \frac{U_0}{4 R_{el}} (\Delta R_2 - \Delta R_1). \quad (447)$$

The resistance changes of radial and tangential strain gauges on a thick circular membrane can be calculated with (83) and (84) on page 56, respectively, and an example is shown in Fig. 40c. Inserting the resistance changes of radial and tangential strain gauges into the above equation for  $R_1$  and  $R_2$ , respectively, yields:

$$U_m = \frac{3}{32} \frac{U_0 \Delta p}{E_M d_M^2} (1 - \nu_M^2) ((1 - 2\nu)r_t^2 + (1 + 2\nu)r_r^2). \quad (448)$$

In the above equation,  $r_t$  and  $r_r$  are the effective distances of tangential and radial strain gauges, respectively, from the center of the membrane. Young's modulus,

thickness of the membrane, pressure drop, Poisson’s ratio of membrane, and strain gauges are denoted by  $E_M$ ,  $d_M$ ,  $\Delta p$ ,  $\nu_M$ , and  $\nu$ , respectively. Equation (448) describes the measured voltage  $U_m$  as a function of the pressure drop  $\Delta p$ , when radial and tangential strain gauges on a thick circular membrane are combined to a half bridge. That is, (448) is the characteristic curve of the pressure sensor. However, the stress of the membrane has not been taken into account for the calculation.

The effect of residual stress can be included in the calculation, if the resistance change is calculated with (89) and (90) (page 57) as a function of membrane deflection  $w_0$  and inserted into (447):

$$U_m = \frac{U_0 d_M w_0}{2 R_M^4} ((1 - 2 \nu)r_t^2 + (1 + 2 \nu)r_r^2). \tag{449}$$

The above equation describes the output voltage of radial and tangential strain gauges from metal on the surface of a thick circular membrane. Equation (449) shows that the output voltage is a linear function of membrane deflection and it is not a function of Young’s modulus, Poisson’s ratio, and residual stress. Each combination of pressure and membrane properties which yields a certain deflection will result in the same output voltage.

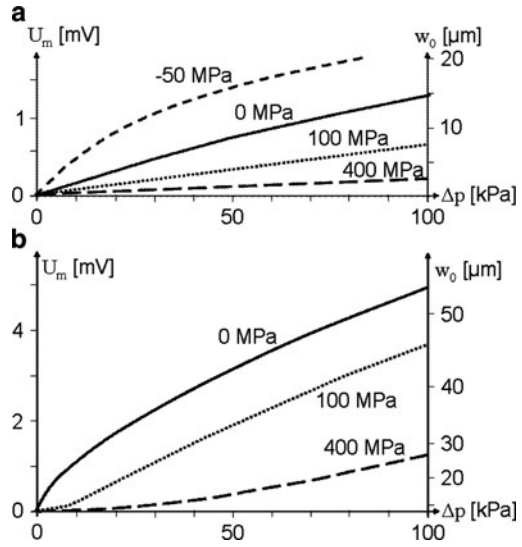
The effect of residual stress on output voltage  $U_m$  can be calculated now by inserting the deflection  $w_0$  as a function of the pressure drop into (449). Since the membrane is thick, its deflection is smaller than its thickness and the third term in the parenthesis of (53) can be neglected. As a result, the output voltage is obtained as a function of the pressure difference  $\Delta p$  and the residual stress  $\sigma_0$  of the membrane:

$$U_m = \frac{1}{8} \frac{(1 - 2 \nu)r_t^2 + (1 + 2 \nu)r_r^2}{4/3(E_M/(1 - \nu_M^2))d_M^2 + R_M^2} \sigma_0 U_0 \Delta p. \tag{450}$$

The above equation is the characteristic curve of the pressure sensor as a function of the residual stress. It shows that the output voltage  $U_m$  is a linear function of the pressure difference. If the deflection gets larger, the characteristic curve of the pressure sensor needs to be calculated in a different way: The output voltage as a function of membrane deflection is calculated with (449), the pressure difference as a function of deflection is calculated with (53), and the output voltage is drawn over the pressure difference. In Fig. 208a, the characteristic curve calculated this way is shown for a membrane with radius, thickness, Young’s modulus, and Poisson’s ratio of 1 mm, 20  $\mu\text{m}$ , 120 GPa, and 0.3, respectively. On the membrane, there are radial and transversal metal strain gauges at effective radii of 600 and 900  $\mu\text{m}$ , respectively, with a Poisson’s ratio of 0.3. The voltage support  $U_0$  is 10 V.

As seen in Fig. 208a, the characteristic curve of a thick membrane is linear, if its deflection is small compared with its thickness (20  $\mu\text{m}$ ). The slope, i.e., the sensitivity of the sensor, is a strong function of residual stress. That is, the reason why

**Fig. 208** Characteristic curves of pressure sensors with (a) a thick and (b) a thin membrane as a function of residual membrane stress  $\sigma_0$ ; and membrane deflection  $w_0$  corresponding to output voltage  $U_m$



pressure sensors are mounted at the free end of a tube or a similar arrangement which avoids that outer forces can change membrane stress (cf. Fig. 207). When membrane deflection gets larger, the characteristic curve is no longer linear, especially when there is some compressive residual stress on the order of magnitude of the critical stress. In the example shown in Fig. 208a, the critical stress calculated with (54) (page 41) is -70 MPa.

If the membrane is deflected more than its thickness, it is called a thin membrane and the resistance change of radial and transversal strain gauges from metal on a circular membrane as a function of deflection can be calculated with (92) and (93) on page 59. Inserting (92) and (93) into (447) yields the output voltage of a pressure sensor employing a circular thin membrane with radius  $R_M$ :

$$U_m = \frac{1}{6} U_0 \frac{w_0^2}{R_M^4} ((1 - \nu)r_t^2 + (1 + \nu)r_r^2). \tag{451}$$

The resistance change of a thin membrane is due to the strain of its neutral fiber. Therefore, the output voltage is a function of the square of membrane deflection. Figure 208b shows the characteristic curve of a pressure sensor with the same parameters as the one in Fig. 208a but with a thickness of just  $d_M = 2 \mu\text{m}$  instead of  $20 \mu\text{m}$ . The pressure difference was calculated with (53) as for Fig. 208a, but the output voltage was obtained from (451). The scale on the right of Fig. 208b shows that the output voltage is a quadratic function of membrane deflection  $w_0$ .

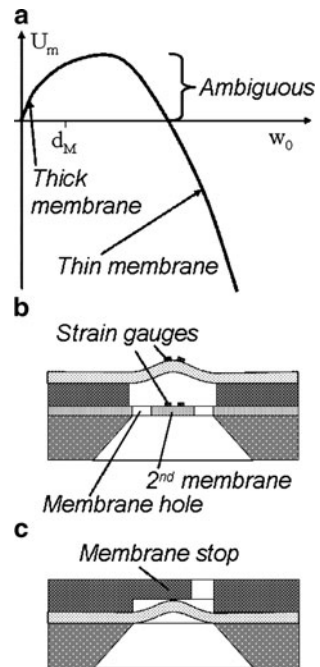
The difference between a thick and a thin membrane is just its deflection. Therefore, a thick membrane becomes a thin one when the pressure drop  $\Delta p$  is enhanced and the deflection  $w_0$  gets larger than membrane thickness  $d_M$ . The resistance change of radial and tangential strain gauges on circular thick and thin

membranes are shown in Fig. 40 (page 56) and 42 (page 59), respectively. The resistance of a radial strain gauge at the rim of a thick membrane becomes smaller with increasing deflection, but it becomes larger when the deflection is larger than membrane thickness and the membrane is called a thin membrane. For a transversal strain gauge, the situation is similar but with the opposite trend as a function of increasing deflection. As a consequence, there is a certain range of deflections which are about as large as the thickness of the membrane where the characteristic curve is ambiguous (cf. Fig. 209a).

To avoid that a pressure sensor designed with a thick membrane enters the ambiguous range, it may be appropriate to mount a membrane stop as shown in Fig. 209c at a distance corresponding to the thickness of the membrane. The electronic of such a sensor can detect an overload when a certain threshold is overcome or when the membrane touches the stop.

If the measurement range shall not be restricted to pressures at which the membrane is clearly thick or thin, there are several ways to avoid the ambiguity of the characteristic curve: Strain gauges in the center, strain gauges at the neutral fiber of the membrane, or deflecting the membrane to the side opposite to the strain gauges.

If strain gauges are arranged at the center of the membrane, the sign of the resistance change does not differ for thick and thin membranes, and, therefore, the characteristic curve will not be linear but unambiguous when the membrane deflection makes the transfer from a thick to a thin membrane. To achieve a compensation



**Fig. 209** Characteristic curve of pressure sensor when membrane deflection exceeds its thickness

for temperature changes affecting the electrical resistance, a second position for conductor paths is required which shows a different resistance change when the membrane is deflected. This may be accomplished by a second membrane next to the first one having one or more large holes and similar conductor paths as the first membrane (cf. Fig. 209b). The conductor paths of both membranes can then be combined to a bridge as shown in Fig. 39 on page 54.

Another way to avoid the ambiguity of the characteristic curve is to place the strain gauges at the neutral fiber inside the membrane. This way, only the strain of the neutral fiber is detected and bending of the membrane will show no effect on the measurement. The output voltage then has to be calculated with (451) also for membrane deflections smaller than membrane thickness. A disadvantage of arranging strain gauges in the neutral fiber of a membrane is that it cannot be distinguished between up and down deflections of the membrane because the output is a function of the square of deflection. This is not important, if the reference pressure is a vacuum but may be a problem when a pressure difference is to be measured.

The fourth (and for most cases the best) solution to avoid ambiguous output signals is to make sure that the membrane is deflected into the direction opposite to the side where the strain gauges are placed. This way, the sign of the resistance change shown in Fig. 40 (page 56) is reversed and the trend of the resistance change of strain gauges at the rim of the membrane does not change with the transition from a thick to a thin membrane.

Most often a square thick silicon membrane is employed for pressure sensors with p-doped strain gauges in a monocrystalline membrane. This is described on page 61f. The strain gauges are orientated parallel to the  $\langle 110 \rangle$ -direction of the silicon lattice. They are arranged parallel and perpendicular to the edges at the rim of the square membrane as shown in Fig. 44 (page 61).

As an approximation, the radial  $\sigma_R$  and tangential  $\sigma_T$  stress of a circular thick membrane with radius  $R_M$  calculated with (79) and (80) (page 55), respectively, are used instead of the stresses perpendicular and parallel to the edge of a square membrane with an edge length  $a_M = 2 R_M$ :

$$\sigma_R = \frac{3}{8} \frac{\Delta p}{d_M^2} \left( \frac{a_M^2}{4} (1 + \nu_M) - r^2 (3 + \nu_M) \right), \quad (452)$$

$$\sigma_T = \frac{3}{8} \frac{\Delta p}{d_M^2} \left( \frac{a_M^2}{4} (1 + \nu_M) - r^2 (1 + 3 \nu_M) \right). \quad (453)$$

In the above equations,  $r$  is the effective distance of the strain gauges from the center of the membrane. Therefore,  $r$  will be a bit less than  $a_M/2$ . The effective distance between membrane center and strain gauges perpendicular  $r_r$  and parallel  $r_l$  to the edges may be different. As a consequence, it needs to be distinguished between the stresses of perpendicular ( $\sigma_{R,r}$  and  $\sigma_{T,r}$ ) and parallel ( $\sigma_{R,l}$  and  $\sigma_{T,l}$ ) strain gauges. Equations (452) and (453) are inserted into (22) on page 23 and the

piezoresistive coefficients of p-doped silicon in Table 1 (page 25) yielding for the resistance change of perpendicular  $\Delta R_{el,R}$  and parallel  $\Delta R_{el,T}$  strain gauges:

$$\frac{\Delta R_{el,R}}{R_{el}} = \pi_l \sigma_{R,r} + \pi_t \sigma_{T,r} + \alpha_T \Delta T. \tag{454}$$

$$\frac{\Delta R_{el,T}}{R_{el}} = \pi_l \sigma_{T,t} + \pi_t \sigma_{R,t} + \alpha_T \Delta T. \tag{455}$$

Inserting the above two equations into (447), yields the output voltage of a half bridge on a square thick membrane from monocrystalline silicon. Usually on such a membrane, there is not a half bride but a full bridge as shown in Fig. 39 on page 54. As a consequence, the output voltage is doubled and the *characteristic curve of a thick square silicon membrane* is:

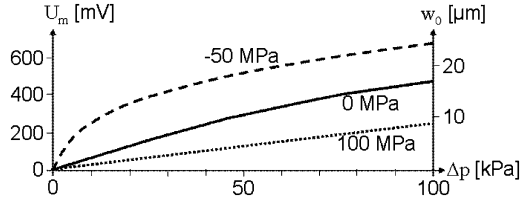
$$U_m = \frac{3}{16} \frac{U_0 \Delta p}{d_M^2} \times (\pi_l [r_r^2(1 + 3 \nu_M) - r_t^2(3 + \nu_M)] - \pi_t [r_r^2(1 + 3 \nu_M) - r_t^2(3 + \nu_M)]). \tag{456}$$

The above equation does not include the effects of residual stress and stress due to straining on the characteristic curve. To include these, first the characteristic curve as a function of membrane deflection  $w_0$  is calculated by replacing the pressure difference  $\Delta p$  in the above equation by the pressure difference necessary to deflect a thick square membrane by  $w_0$  (found in Table 2 on page 50):

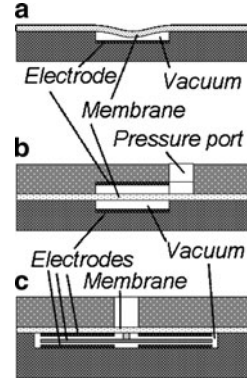
$$U_m = \frac{99}{8} \frac{U_0 E_M d_M w_0}{a_M^4} \times \frac{\pi_l [r_r^2(1 + 3 \nu_M) - r_t^2(3 + \nu_M)] - \pi_t [r_r^2(1 + 3 \nu_M) - r_t^2(3 + \nu_M)]}{1 - \nu_M^2}. \tag{457}$$

Then, the output voltage calculated with the above equation for certain values of the deflection is plotted over the pressure difference calculated for the same deflections as calculated with (67) on page 46 which is also found in Table 2. The characteristic curve of a silicon pressure sensor calculated with (67) and (457) is shown in Fig. 210. The same parameters were used as for Fig. 208a with the exception of the effective radius of the radial strain gauges which was assumed to be 800  $\mu\text{m}$  instead of 600  $\mu\text{m}$ . Figure 210 shows that the characteristic curve of a silicon pressure sensor is similar for metal strain gauges on a membrane, but the output signal of the silicon sensor is much larger because of the larger gauge factors of monocrystalline silicon. The output of a real silicon sensor will be approximately 20% less as shown in Fig. 210 because, for the calculation, the same Young’s modulus (120 GPa) and Poisson’s ratio (0.3) were used as for Fig. 208 to allow for a better comparison.

**Fig. 210** Characteristic curves of a silicon pressure sensor as a function of membrane stress



**Fig. 211** Capacitive pressure sensors



In *capacitive pressure sensors*, the deflection of the membrane is measured as the capacity change between the membrane and a counter electrode as shown in Fig. 211a. Similar as for strain gauges, it is advantageous to arrange the capacitor between membrane and counter electrode in a bridge to measure small capacitance changes starting from 0 V. Different than in the case of strain gauges, the bridge is not needed to compensate for temperature changes because the electrostatic force is not a function of temperature. The other electronic elements in the bridge do not need to be capacitors but also may be ohmic resistances, if their resistance is not affected too much by temperature changes. In any case, the output is an alternating voltage with an amplitude which is a function of membrane deflection and pressure difference.

The capacitance between two flat electrodes is given in (221) on page 133. Obviously, the capacitance is not a linear function of the distance between the electrodes. Besides this, the deflected membrane is not flat, and, as a consequence, the electric field is concentrated at the membrane center. Therefore, no linear characteristic curve can be achieved.

An interesting alternative is a *balanced pressure sensor*. In such a sensor, the pressure force is compensated by an electrostatic force. The principal design of such a sensor is shown in Fig. 211b. A voltage is applied between the membrane and the electrode above it, and this voltage is controlled such that the membrane is held in its idle state. The voltage  $U_m$  necessary to hold the membrane then is a measure of the pressure difference  $\Delta p$ . The position of the membrane can be

determined either by strain gauges or by the capacity between membrane and another electrode, e.g., below it.

If the membrane is held in its idle state, electrostatic force  $F_C$  and pressure need to be equal. The electrostatic force acting on a flat membrane at a distance  $d_C$  to an electrode with area  $A_C$  is calculated with (222) (on page 131):

$$F_C = \frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C^2} U_m^2 = \Delta p A_M \Rightarrow U_m = d_C \sqrt{2 \frac{\Delta p A_M}{\epsilon_0 \epsilon_r A_C}}. \tag{458}$$

The above equation shows that the characteristic curve of a pressure sensor compensating the pressure by an electrostatic force is a root function (cf. Fig. 212). It also shows that it is only a function of the geometry of electrodes and membrane but not a function of the elastic properties of the membrane and its thickness. Since the dimensions of the sensor only vary very little with temperature, the characteristic curve is only a weak function of temperature.

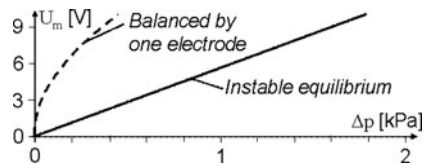
It needs to be noted that comparatively large voltages are required to compensate usual pressure differences. If very small pressure differences are to be measured, this is an advantage because the sensor is very sensitive. If the measurement range shall be extended to larger pressures, the design can be changed as shown in Fig. 211c and the ratio of the areas of membrane  $A_M$  and electrodes  $A_C$  can be reduced [74].

In principle, a linear characteristic curve can also be achieved. To the upper electrode in Fig. 211b or c, a DC voltage of  $+U_0$  can be applied and to the lower electrode a voltage  $-U_0$ . If there is no pressure difference and the membrane is connected to earth potential, the membrane is attracted to both electrodes by the same force and it is in an instable equilibrium of forces between the electrodes. When some pressure difference is acting on the membrane, a voltage  $\Delta U$  has to be applied to the membrane to hold it in its idle state. The membrane is lifted up by the voltage  $U_0 - \Delta U$  and pulled down by  $U_0 + \Delta U$ . The resulting force is equal to the force generated by the pressure difference and can be calculated according to (222) (page 131):

$$F_C = \frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C^2} [(U_0 - \Delta U)^2 - (U_0 + \Delta U)^2] = 2 \epsilon_0 \epsilon_r \frac{A_C}{d_C^2} U_0 \Delta U = A_M \Delta p \tag{459}$$

$$\Rightarrow \Delta U = \frac{A_M}{A_C} \frac{d_C^2}{2 \epsilon_0 \epsilon_r U_0} \Delta p. \tag{460}$$

**Fig. 212** Characteristic curves of pressure sensors balanced by an electrostatic force





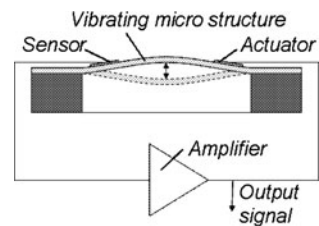
The voltage  $\Delta U$  applied to the membrane now is the measure of the pressure difference and the equation above is the characteristic curve of the sensor. The characteristic curve is linear and it is only a function of the ratio of the areas of membrane and electrodes, their distance  $d_C$ , and the voltage supplied to the stationary electrodes  $U_0$ . Balanced pressure sensors are similar to the balanced acceleration sensors described on page 305f. The pressure force  $\Delta p A_M$  in (458) and (460) is just replaced by the acceleration force  $a_m m_0$  in (490) (page 305) and (491) (page 306), respectively. In Fig. 212, there are shown the characteristic curves of pressure sensors balanced by one electrode and in an instable equilibrium. A distance of 1  $\mu\text{m}$  was assumed between membrane and electrodes and a driving voltage of 10 V in the case of the instable equilibrium.

The pressure range which can be measured is limited by the maximum voltage which can be supplied to the electrodes. It is reasonable assuming the voltage  $\Delta U$  applied to the membrane not larger than  $U_0$  applied to the electrodes. In Fig. 212, 10 V correspond to a maximum pressure of 450 and 1,780 Pa. Of course, it needs to be made sure that the membrane needs to be thin and large enough to show an easily detectable deflection when a small pressure is applied.

The deflection of the membrane of a pressure sensor can also be measured by a vibrating microstructure whose frequency is a measure of the pressure difference. Sensors with a frequency output are called *frequency analog sensors*. The advantage of a frequency as the output signal is that it is hardly affected by electromagnetic noise.

On the other hand, there is a mechanism required which excites a microstructure to vibrations which are a function of the measurand. Typically, the resonance frequency of a beam clamped at both ends is employed. Heating such a beam (cf. page 164), driving it as a piezo (cf. page 144), or electrostatically attracting it (cf. page 131) are methods to excite vibrations. In Fig. 213, it is shown how, in principle, resonance vibrations can be excited and observed. Typically, the deflection of a microstructure is excited by some actuator and from its vibrations a signal is derived with a sensor. This sensor signal is amplified and employed to drive the actuator again. After a few deflections, the microstructure is vibrating at its resonance frequency. As a sensor either strain gauges, the change of the capacity between the vibrating beam and a counter electrode, or the piezoelectric effect can be employed.

A vibrating beam can be etched into the surface of the membrane of a pressure sensor (cf. Fig. E3 on page 16) [14]. The membrane is deflected by a pressure difference and the strain at its surface is transferred to the beam. The fundamental



**Fig. 213** Principle of driving a frequency analog sensor

frequency  $f_1$  of a beam as a function of its stress  $\sigma$  is described by the equation found in Table 10 on page 102:

$$f_1 = \frac{11.2}{\pi L_B} \sqrt{\frac{E_B I}{L_B m_B} + \frac{b_B d_B L_B}{4 \pi^2}} \sigma. \quad (461)$$

The strain  $\varepsilon$  of a circular membrane at its center is obtained from (81) and (82) on page 55 and 56. It can also be employed as an approximation for the stress of a square membrane with an edge length  $a_M = 2 R_M$ :

$$\varepsilon = \frac{3 \Delta p R_M^2}{8 E_M d_M^2} (1 - \nu_M^2) \Rightarrow \sigma_D = \varepsilon E_B = \frac{3 E_B R_M^2}{8 E_M d_M^2} (1 - \nu_M^2) \Delta p. \quad (462)$$

In the above equation, the stress  $\sigma_D$  due to straining of the beam is calculated by Hooke's law. It differs from the stress of the membrane [calculated with (79) and (80) on page 55] by the power of Poisson's ratio even if beam and membrane have the same Young's modulus, because the beam is strained only in longitudinal direction.

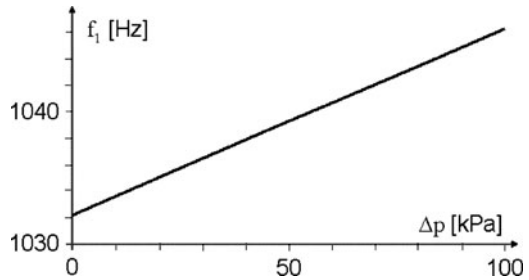
Inserting (462) into (461) and taking into account that there may have been some residual stress  $\sigma_0$  in the beam when the membrane is not deflected yields:

$$f_1 = \frac{11.2}{\pi L_B} \sqrt{\frac{E_B I}{L_B m_B} + \frac{b_B d_B L_B}{4 \pi^2} \left( \sigma_0 + \frac{3 E_B R_M^2}{8 E_M d_M^2} (1 - \nu_M^2) \Delta p \right)}. \quad (463)$$

A difficulty of frequency analog sensors is that the measured frequency could not be the desired resonance of the beam but another resonance, e.g., of the membrane or another part of the sensor. It is also necessary to make sure that the beam itself does not vibrate in an undesired mode, e.g., perpendicular to the expected direction or in a higher resonance. The frequency may not be altered by particles attached to the beam or by a fluid surrounding it. Therefore, it is desirable to mount the beam inside of the reference vacuum of the sensor.

If the dimensions are chosen such that the bending moments are dominating the frequency [In (463) the first term in the square root is much larger than the second one.], the characteristic curve is nearly linear. The ratio of membrane radius to thickness should be large to make the second term in the parenthesis large compared with the residual stress and insensitive to changes of that stress. In Fig. 214, the characteristic curve calculated with (463) is shown for a sensor equipped with a membrane, 20  $\mu\text{m}$  and 2 cm in thickness and radius, respectively, and a beam, 4 mm, 2 mm, and 2  $\mu\text{m}$  in length, width, and thickness, respectively. Membrane and beam are assumed to be fabricated from silicon ( $E_B = 150 \text{ GPa}$ ,  $\nu = 0.3$ ,  $\rho = 2,330 \text{ kg/m}^3$ ). The sensor is based on the ability to measure the resonance frequency with high precision. If the frequency is not measured by counting the number of vibrations in a certain time but by measuring the time between two zero crossings, a remarkable reaction time on the order of 1 ms can be achieved.

**Fig. 214** Characteristic curve of a frequency analog pressure sensor



In principle, the membrane itself could be employed as the resonating microstructure. However, it cannot be avoided that the membrane is in contact to the fluid whose pressure is to be measured. Therefore, the characteristic curve would be a function of the properties of the fluid, which contributes to the vibrating mass, may contain particles which stick to the membrane and enhance the vibrating mass, and damp membrane vibrations by its viscosity and sound emission (cf. page 92f.).

Certainly, it is also very important to make sure that outer forces or temperature changes do not alter the residual stress of the vibrating microstructure. There always will remain an effect of temperature changes on Young's modulus and the dimensions of the microstructure which needs to be acceptable or compensated.

## Exercises

### Problem 49

- (a) Which output voltage is expected for a pressure sensor deflected by a pressure difference of 100 kPa? The sensor is equipped with radial and tangential strain gauges on a circular thick membrane. Properties and dimensions of the sensor are shown in the following table:

Voltage supply	5 V
Thickness of the membrane	5 $\mu\text{m}$
Radius of the membrane	1 mm
Young's modulus of the membrane	120 GPa
Poisson's ratio of the membrane	0.3
Poisson's ratio of strain gauges	0.3
Effective radius of tangential conductor paths	900 $\mu\text{m}$
Effective radius of radial conductor paths	700 $\mu\text{m}$

- (b) By how many percent is the output signal of the pressure sensor changed, if the membrane gets a residual stress of 100 MPa?

**Problem 50**

A pressure sensor is equipped with a thin circular membrane. On the membrane, there are radial and tangential strain gauges from gold. Properties and dimensions are given in the following table. The two strain gauges on the membrane are combined to a half bridge as shown in Fig. 39 on page 54. The potentiometer is a part of the electronics which is not on the membrane.

Effective radius of tangential strain gauge	4 mm	Young's modulus of membrane	120 GPa
Thickness of strain gauges	100 nm	Effective radius of radial conductor paths	4.5 mm
Width of strain gauges	5 $\mu\text{m}$	Radius of membrane	5 mm
Poisson's ratio of membrane	0.3	Membrane thickness	2 $\mu\text{m}$
Poisson's ratio of strain gauges	0.3	Specific electrical resistance of gold	$22 \times 10^{-9} \Omega \text{ m}$
Length of each strain gauge	25 mm	Residual stress of membrane	150 MPa

- (a) What is the maximum supply voltage, if power consumption of the sensor shall be limited to 1 mW?
- (b) Calculate the characteristic curve between 0 and 100 kPa, if the supply voltage is 10 V.

Hint: Employing Excel or a similar code will help very much.

- (c) What is the smallest pressure which can be measured, if the signal needs to be at least 1 mV to be distinguished from noise?
- (d) Now instead of the sensor described above, an electrostatic pressure sensor with pressure compensation shall be investigated. The distance between membrane and counter electrode is 1  $\mu\text{m}$  ( $\epsilon_0 = 8.9 \times 10^{-12} \text{ A s}/(\text{V m})$ ,  $\epsilon_r = 1$ ). What is the minimum pressure required for an output voltage of 1 mV?

**Problem 51**

The pressure sensor shown in Fig. E3 on page 16 is a so-called frequency analog sensor. That is, its output signal is a frequency. Which frequency is expected at a pressure of 1 Pa and 100 kPa, respectively?

Thickness of beam	20 $\mu\text{m}$	Poisson's ratio of membrane	0.45
Width of beam	100 $\mu\text{m}$	Density of beam	2,330 $\text{kg}/\text{m}^3$
Length of beam	500 $\mu\text{m}$	Radius of membrane	2 mm
Young's modulus of beam and membrane	100 GPa	Thickness of membrane	100 $\mu\text{m}$

# Flow Sensors

The measurement of volume flow or flow velocity has a lot of applications in modern life. For example, for the control of the engine of a car with low exhaust emission, it is essential to know the fuel and air flow to the engine. In industry, it is necessary to apply well-defined amounts of lubricant into the bearing of gear wheels of watches or to dose glue onto small areas next to positions where no bond is desired. It is evident that, in chemical and pharmaceutical industry, mixing of substances in the correct ratio is important.

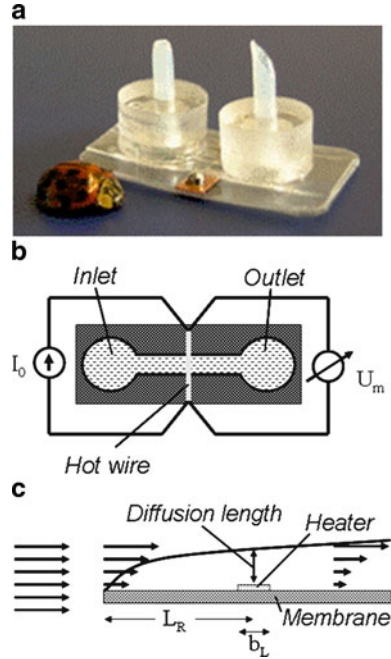
There is a variety of measurement principles in use and several ways are known how microsensors for flow meters can be built. In the following, the most important principles which can be employed in microsensors are introduced, i.e., thermal flow sensors and the measurement of the pressure difference over a capillary.

Maybe the simplest flow sensor is just a thin wire heated by an electrical current. When a fluid is passing the wire, it takes away a part of the heat and lowers the temperature of the wire. Thus, the temperature of the wire is a measure of the flow velocity. The temperature of the wire can be determined from its electrical resistance, which is nearly a linear function of temperature [cf. (18) and (19), page 22].

Figure 215a, b shows a micro-flow sensor employing such a wire. This kind of flow sensors is called hot-wire anemometers (Anemos is the Greek word for wind.). In Fig. 215b, a constant current is going through the wire. If the temperature of the wire is lowered by the flow, its resistance is decreased and the voltage measured over the wire becomes smaller.

Obviously, the reaction time of the sensor gets short, if the temperature of the wire is quickly changed by the flow. Therefore, the wire should have a small volume and a large surface area. In microtechnique, this is achieved with a thin conductor path on a thin carrier membrane as shown in Fig. 215c. The velocity of the fluid develops a parabolic profile when it gets in contact to the membrane. Directly on the surface of the membrane and the heater, the flow velocity is zero. Therefore, no heat could be taken away from the heater, if there would be no heat transfer by diffusion. As an approximation, it can be assumed that over a certain diffusion length there would be no heat transport by convection, but only by diffusion and behind that length the heat is taken away by the fluid.

**Fig. 215** Photo of a micro-hot-wire anemometer (a), a schematic top view (b), and cross-section of a heater wire on a membrane (c)



As a consequence of the development of the parabolic profile, the diffusion length gets the longer, the larger the distance  $L_R$  is from the edge of the membrane to the center of the heater. The heat  $Q_K$  transported away by the flow is a function of the thermal conductivity  $\lambda_F$ , the viscosity  $\eta$ , and density  $\rho_F$  of the fluid, the temperature difference  $\Delta T$  between hot wire and fluid, the area  $A_S$  of the heater wire, and the mean flow velocity  $v$  far away from the membrane.  $Q_K$  is estimated with King's law [71]:

$$Q_K = a \lambda_F A_S \Delta T \sqrt{\frac{\rho_F v}{L_R \eta}} \tag{464}$$

King's law shows that the heat taken away from the heater is a function of the square root of the flow velocity. It was derived for laminar flows not effected by nearby channel walls [71]. Obviously, in microtechnique this is not fulfilled. The constant (a) up to now cannot be calculated analytically in any way but needs to be measured during calibration of the sensor.

Equation (464) actually is applicable for a conductor path which is narrow compared with its distance  $L_R$  to the rim. However, the designer will try to make  $L_R$  as small as possible to achieve a sensitive sensor. Therefore, the heat transported away from the sensor needs to be calculated from the following integral:

$$\begin{aligned}
 Q_K &= \int_{L_R - (b_L/2)}^{L_R + (b_L/2)} a \lambda_F L_L \Delta T \sqrt{\frac{\rho_F v}{b \eta}} db \\
 &= a \lambda_F L_L \Delta T \sqrt{\frac{\rho_F v}{\eta}} 2 \left( \sqrt{L_R + \frac{b_L}{2}} - \sqrt{L_R - \frac{b_L}{2}} \right). \quad (465)
 \end{aligned}$$

The area  $A_S$  of the heater wire in the above equation is calculated from the product of its length  $L_L$  (perpendicular to the flow direction) and its width  $b_L$  (in flow direction). If the heater is not a flat conductor path but a cylindrical wire with radius  $R_B$ , it appears to be a good approximation to use half of the circumference of the wire for the width  $b_L = \pi R_B$  and the distance to the rim is  $L_R = \frac{1}{2} \pi R_B$ . If the heater is mounted in the center of a microchannel, it is cooled by the flow from both sides, and therefore, the area is doubled and a factor of two has to be multiplied to the equation:

$$Q_K = 4 a \lambda_F L_L \Delta T \sqrt{\frac{\pi R_B \rho_F v}{\eta}} = 4 a \lambda_F L_L \Delta T \left( \frac{\pi R_B \rho_F v}{\eta} \right)^{0.5}. \quad (466)$$

Some measurements show that the exponent in the above equation can differ from 0.5 and tends to be near 0.3.

For a heater mounted next to the rim of a supporting membrane ( $L_R = b_L/2$ ), the difference between (465) and (464) is 41%. However, this is normally not discovered because the constant (a) cannot be obtained from calculation. The constant (a) is a function of the position of the heater in the channel and the flow profile around the heater.

King's law only describes the heat taken off the heater by the flow of the fluid. Much more heat is taken away by heat conduction of the heater towards the sidewalls of the channel. Besides this, there is also the heat conductivity of the fluid which will transport heat to the side walls even if there is no flow. The heat taken away by heat conduction is entered now into (465) by the quantity  $\Lambda_e$  which describes how much heat power per temperature difference is conducted off the wire. The unit of  $\Lambda_e$  is W/K:

$$Q_K = 2 a \lambda_F L_L \Delta T \sqrt{\frac{\rho_F v}{\eta}} \left( \sqrt{L_R + \frac{b_L}{2}} - \sqrt{L_R - \frac{b_L}{2}} \right) + \Lambda_e \Delta T. \quad (467)$$

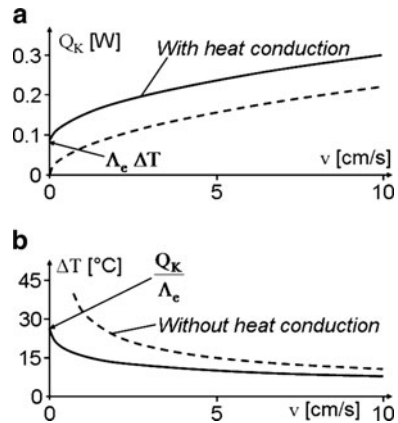
There are two principle ways how to drive an anemometer. One possibility is to drive the wire with constant electrical power (This is similar to constant voltage or constant current.) and to detect the temperature of the wire as a measure of the flow. As a result of this, the heat flow  $Q_K$  in (467) is constant and the temperature difference  $\Delta T$  is calculated as:

$$\Delta T = \frac{Q_K}{2 a \lambda_F L_L \sqrt{\frac{\rho_F v}{\eta}} (\sqrt{L_R + (b_L/2)} - \sqrt{L_R - (b_L/2)}) + \Lambda_e} \quad (468)$$

The other possibility to drive an anemometer is to employ an electronic circuit which keeps the temperature (resistance) of the wire constant by adapting the power. In this case, the electrical power which is equal to the heat flow  $Q_K$  is a measure of the flow and can be calculated with (467). Figure 216a, b shows the characteristic curves of an anemometer with a distance of the heater to the edge of the membrane of 10  $\mu\text{m}$ , a heater length  $L_L$  and width  $b_L$  of 0.5 mm and 0.1 mm, respectively, and water as the fluid [ $\eta = 1 \text{ mPa s}$ ,  $\rho_F = 1,000 \text{ kg/m}^3$ ,  $\lambda_F = 0.6 \text{ W/(m K)}$ ] calculated with (467) at a temperature difference of 10°C and with (468) at a heating power of 234 mW, respectively. The heater is assumed to be in the center of the channel and an extra factor of 2 was applied in both equations. For  $\Lambda_e$  and the constant (a), 8 mW/K and 8 were used, respectively.

In Fig. 216a, b, there are also shown the calculations without heat conduction to demonstrate that it is important to take it into account. If heat conduction is reduced by a suitable design, e.g., employing a thin supporting membrane with small heat conductivity, the sensor is more sensitive.

The electronics required to drive a hot-wire anemometer at constant current is comparatively simple. On the other hand, when the anemometer is driven at constant temperature, its reaction time is much shorter because it is not necessary to wait for the temperature change. The reaction time with temperature change can be calculated from the power balance at the wire. The electrical power  $P_{el}$  input to the wire goes to the fluid according to King’s law, to the walls of the channel due to heat conductivity, and to the temperature change of the wire (To ease the following calculation, it is assumed that there is no distance of the heater to the rim of the carrier membrane and that it is mounted in the center of the flow channel. The calculation can easily be extended to the general case.):



**Fig. 216** King’s law (a) and temperature difference of a hot-wire anemometer driven at constant electrical power (b) calculated with (467) and (468)



$$P_{el} = 4 a \lambda_F L_L \Delta T \sqrt{\frac{b_L \rho_F v}{\eta}} + \Lambda_e \Delta T + C_D \frac{dT}{dt}. \quad (469)$$

The first term on the right side of the above equation is the heat loss of the wire due to King's law. The second term is the heat conducted to the wall. The third term is the power required to change the temperature of the wire and its carrier membrane which together have a heat capacity  $C_D$ .

Equation (469) is a differential equation which can be solved with the following ansatz which contains the reaction time of the anemometer  $\tau_A$ , the temperature difference between wire and fluid before  $\Delta T_a$  and a long time after  $\Delta T_e$  the velocity of the fluid has changed:

$$\Delta T = \Delta T_e + (\Delta T_a - \Delta T_e)e^{-t/\tau_A}. \quad (470)$$

When the flow velocity and the electrical power are constant, after some time the temperature difference between wire and fluid is also constant and the last term in (469) is zero, resulting in:

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v}{\eta}} + \Lambda_e \right) \Delta T. \quad (471)$$

The above equation needs to be fulfilled both before (index of velocity and temperature difference is a) and after (index e) a change of the mean flow velocity:

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v_a}{\eta}} + \Lambda_e \right) \Delta T_a = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v_e}{\eta}} + \Lambda_e \right) \Delta T_e \quad (472)$$

$$\Rightarrow \Delta T_a = \frac{P_{el}}{4 a \lambda_F L_L \sqrt{(b_L \rho_F v_a/\eta)} + \Lambda_e} \quad \text{and} \quad \Delta T_e = \frac{P_{el}}{4 a \lambda_F L_L \sqrt{(b_L \rho_F v_e/\eta)} + \Lambda_e}. \quad (473)$$

Now the ansatz (470) is inserted into the differential equation (469):

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v_e}{\eta}} + \Lambda_e \right) \left( \Delta T_e + [\Delta T_a - \Delta T_e]e^{-t/\tau_A} \right) - \frac{C_D}{\tau_A} [\Delta T_a - \Delta T_e]e^{-t/\tau_A}. \quad (474)$$

At the time  $t = 0$ , when the velocity change occurs, the above equation yields:

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v_e}{\eta}} + \Lambda_e \right) \Delta T_a - \frac{C_D}{\tau_A} [\Delta T_a - \Delta T_e]. \quad (475)$$

From the above equation, the reaction time can be calculated. With (473), it can be expressed in different ways:

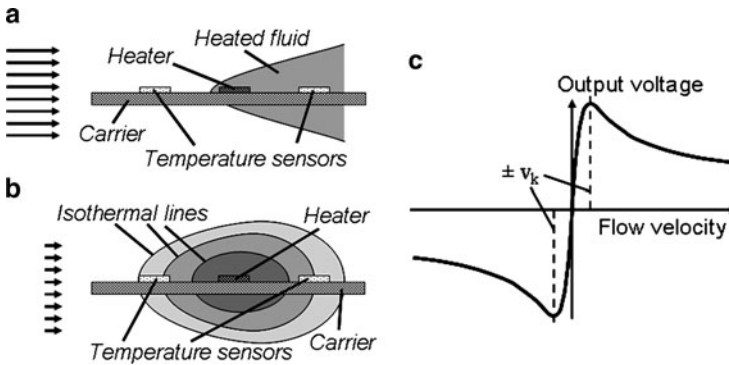
$$\begin{aligned}\tau_A &= \frac{C_D[\Delta T_a - \Delta T_e]}{(4 a \lambda_F L_L \sqrt{(b_L \rho_F v_e / \eta)} + \Lambda_e) \Delta T_a - P_{el}} \\ &= \frac{C_D}{4 a \lambda_F L_L \sqrt{(b_L \rho_F v_e / \eta)} + \Lambda_e} = \frac{C_D}{P_{el}} \Delta T_e.\end{aligned}\quad (476)$$

It can be concluded that for a short reaction time  $\tau_A$  of the sensor the heat capacity  $C_D$  of the heater and its carrier should be small, the area of the heater should be large, the distance  $L_R$  of the heater to the edge of the carrier membrane should be short, and the sensor should be placed where the flow velocity is large.

Large heat conductivity to the channel walls reduces the reaction time. However, the sensitivity of the sensor is also reduced by large heat conductivity and the power consumption of the sensor is increasing. As a consequence, the carrier and the heater should be designed as thin as possible and the connection to the channel walls should be narrow and long. The distance of the heater to the walls should be as large as possible, i.e., in the center of the flow channel. Making the channel wider at the position of the sensor does not appear to solve the problem because the flow velocity is reduced that way.

The hot-wire anemometer has a simple design and can achieve short reaction times on the order of milliseconds. However, there are also disadvantages: It is not possible to detect the direction of the flow, the characteristic curve of the sensor is not linear and it is a function of temperature, density, heat conductivity, and viscosity of the fluid. Therefore, reliable measurements are only possible when the properties of the fluid are known and the temperature is kept constant. Besides this, gas bubbles in a liquid or droplets in a gas may disturb the measurement very much or even can destroy the heater when cooling suddenly has to be provided only by a gas bubble.

The direction of flow can be detected when two more wires are added to the heater as shown in Fig. 217a. The heated fluid is transported by the moving fluid to one of the additional wires which serve as temperature sensors. The electrical resistance of the sensing wires is measured with a small electrical current not significantly heating up these wires. The temperature (resistance) difference of the two conductor paths neighboring the heater is employed as a measure of the flow velocity. The temperature difference of the two sensing wires is no longer a function of the temperature of the fluid, and the warmer wire indicates the direction of the flow. When both sensing wires are combined in one branch of a Wheatstone bridge, the sign of the output signal indicates the direction of the flow. The output signal is much smaller than the one of an anemometer, but the characteristic curve is similar as the one shown in Fig. 216 because the heat taken away from the heater partly arrives at the downstream sensing wire and the upstream wire is not influenced. The response time is longer because it takes some time for the heat to travel with the fluid to the sensing wire and to heat it up.



**Fig. 217** Hot-wire anemometer (a), calorimeter (b), and characteristic curve of a calorimeter (c)

When the heat transport by the flow is smaller than by heat diffusion, the situation changes and the flow sensor is no longer called an anemometer but a *calorimeter*. The heat is mainly transported away from the heater by the heat conductivity of the fluid and the temperature distribution around the heater is only altered by the flow as indicated in Fig. 217b. The characteristic curve of a calorimeter (cf. Fig. 217c) is nearly linear up to a critical flow velocity  $v_k$  at which the heat transport by the flow is quicker than by diffusion. When the critical flow velocity is exceeded, the flow sensor works as an anemometer and shows a characteristic curve similar as in Fig. 216b.

The critical flow velocity is a function of the distance  $L$  between sensing wire and heater, heat conductivity  $\lambda_F$ , specific heat capacity  $c_f$ , and density  $\rho_F$  of the fluid. It can be estimated by the following equation [72]:

$$v_k = \frac{\lambda_F}{\rho_F c_f L} \Rightarrow L_k = \frac{\lambda_F}{\rho_F c_f v} \tag{477}$$

As shown in the above equation the transition between hot-wire anemometer and calorimeter can also be interpreted as a critical distance  $L_k$  between sensing wires and heater. When the distance is larger, the sensor works as an anemometer, otherwise as a calorimeter.

Obviously, every sensor has two ranges of flow measurement. If it is not known in which range the sensor is working, the output signal is ambiguous. To solve that problem, it is possible to drive the heater as an anemometer and to measure, in addition, the temperature difference between two sensing wires upstream and downstream. Another possibility is to mount more than one pair of sensing wires and to compare the results of all wire pairs. The results which coincide are the right ones.

A common disadvantage of all the principles discussed so far is that the characteristic curves are a function of the properties of the fluid such as viscosity, density, heat capacity, and heat conductivity. As a consequence, a calibration is

required for each kind of fluid, and it is not possible to obtain a reliable flow measurement of an unknown fluid. This disadvantage can be partly avoided by a *flow time measurement* [75]. The heater injects only short heat pulses into the flow and the arrival time of the pulses is measured with sensing wires. Heat conductivity and parabolic flow profile distribute the heat pulse over a longer time, but the temperature peak travels with the speed of the fluid. Only the time required for heating up the fluid by the heater and the sensing wire by the fluid is a function of the properties of the fluid. Therefore, the flow time between heater and sensing wire need to be large compared with the time required for heating or cooling the wire (on the order of ms).

To estimate the time for a temperature change of the heating wire when the electrical current is switched on, (469) is employed again. Since this is the same differential equation as before, the ansatz 470 can be used again. Contrary to the calculation of the response time of a calorimeter, the flow velocity does not change now but the heating power is switched on at  $t = 0$ . As a consequence, the boundary conditions are changed: Before switching on the electrical power, the temperature of the heater and the fluid are the same resulting in  $\Delta T_a = 0$ ; and the temperature difference  $\Delta T_e$  a long time after switching on the power is according to (471):

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v}{\eta}} + \Lambda_e \right) \Delta T_e$$

$$\Rightarrow \Delta T_e = \frac{P_{el}}{4 a \lambda_F L_L \sqrt{(b_L \rho_F v/\eta)} + \Lambda_e}. \quad (478)$$

Inserting  $\Delta T_a = 0$  into (470) yields:

$$\Delta T = \Delta T_e \left( 1 - e^{-t/\tau_A} \right), \quad (479)$$

inserting the above equation into (469) results in

$$P_{el} = \left( 4 a \lambda_F L_L \sqrt{\frac{b_L \rho_F v}{\eta}} + \Lambda_e \right) \Delta T_e \left( 1 - e^{-t/\tau_A} \right) + \frac{C_D \Delta T_e}{\tau_A} e^{-t/\tau_A}, \quad (480)$$

and at  $t = 0$  it is obtained with (478):

$$P_{el} = \frac{C_D \Delta T_e}{\tau_A} \Rightarrow \tau_A = \frac{C_D \Delta T_e}{P_{el}} = \frac{C_D}{4 a \lambda_F L_L \sqrt{(b_L \rho_F v/\eta)} + \Lambda_e}. \quad (481)$$

The above equation is only an estimate of the time required to heat up the fluid to a constant temperature. The heating time should be approximately three times larger than  $\tau_A$  to achieve the optimum fluid temperature. The time required to heat up the sensing wire will also be on the order of  $3 \tau_A$ . Really, the time required

for sensing has to be calculated a bit different, because on the one hand only the wire and not the fluid has to be heated and on the other hand the heat capacity of the fluid will show some influence, if it is not much larger than the one of the wire, because the fluid will get cooler while the wire warms up.  $\tau_A$  is a function of the properties of the fluid and the wire; therefore, the distance between heater and sensing wire needs to be much larger than  $6 \tau_A$  times the mean flow velocity to obtain a result which is not a function of the properties of the fluid. This is the case even for comparatively fast flows because the critical distance  $L_k$  [(cf. (477))] is required between heater and sensing wire making sure that the flow time is measured and not the time of heat conduction.  $L_k$  typically is much larger than the distance the fluid flows in the short time of heating up.

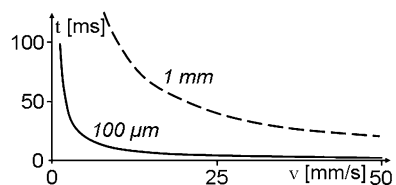
On the other hand, the time of arrival of the temperature peak can be measured with more precision in the near of the heater where the heat pulse is not jet distributed so much. As a result, the sensing wire should be placed not far from the critical distance expected for the slowest flow.

The characteristic curve of a flow time measurement is easily calculated. If the distance of a sensing wire to the heater is  $L$  and the flow velocity in the near of the wire is  $v$ , the flow time  $t$  is given by:

$$t = \frac{L}{v} \tag{482}$$

The characteristic curve of a flow time sensor is shown in Fig. 218. A short reaction time of the sensor can be achieved when the distance  $L$  between heater and sensing wire is short. However, for larger flows, the slope of the characteristic curve is very small for small  $L$  and the sensitivity of the sensor is low. Therefore, a combination of several wires appears to be a good choice allowing for both a short response time for small flows and high accuracy for large flows.

If not the traveling time of the heated fluid from the heater to a sensing wire is measured, but the traveling time from one sensing wire to another one, the time required for heating the fluid cannot affect the measurement. Besides this, the heating up of the two sensing wires should require the same time. Thus, it could be expected that the flow could be measured without any knowledge of the properties of the fluid. Measurements, however, show that this is not true. The measured flow times are a function of the kind of fluid. A possible explanation of this observation is that the viscosity of the fluid is a function of temperature and the flow profile in the center of the channel where the fluid is heated is changed.



**Fig. 218** Characteristic curve of a flow time sensor for different distances  $L$  between heater and sensing wire

Although the flow time measurement is a function of fluid parameters, it is much less sensitive on these parameters than anemometers and calorimeters. Therefore, it is an option to employ a flow time measurement from time to time to recalibrate a sensor anyway consisting of several wires. Disadvantages of flow time measurement are that the reaction time is very large and that the slope of the characteristic curve is very small at large flow velocities. The heat capacity of the fluid needs to be large enough to heat sensing wires.

Flow time measurement is only a weak function of the properties of the fluid. However, it does not work with gasses because their heat capacity is very small and the transported heat is not sufficient to heat up enough the sensing wire. There are also applications such as dosing of drugs for which heating up the fluid is not desirable. An alternative which avoids such problems is the measurement of the *volume displaced by the flow*.

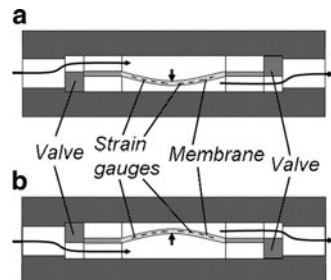
The volume displacement can be measured with a device such as shown in Fig. 219 [76]. The flow is entering into a chamber limited by a membrane. The membrane is deflected by the flow and, that way, displaces the fluid on the side opposite to the entering flow. Thus, the flow entering the device is the same as the flow leaving on the opposite side. When the membrane is approaching the chamber wall, valves at the inlet and outlet of the chamber are switched such that the direction of membrane movement is reversed.

Strain gauges in or on the membrane record its deflection, which is a linear function of the displaced volume [(303), page 170], if the deflection is small compared with the thickness of the membrane (cf. page 38f). When strain gauges from metal are placed on the surface of the membrane, the output of a half bridge, including radial and tangential strain gauges, is described by (449) (page 277) as a function of membrane deflection  $w_0$ . The output voltage as a function of the displaced volume  $V_W$  is obtained by inserting the deflection from (303) into (449):

$$U_m = \frac{U_0 d_M V_W}{\pi R_M^6} ((1 - 2 \nu)r_t^2 + (1 + 2 \nu)r_r^2). \quad (483)$$

When the strain gauges are placed in the neutral fiber (cf. page 6) of the membrane, they measure its strain and not the bending. As a consequence, the

**Fig. 219** Measurement of the volume displaced by the flow [76]. The flow is deflecting a membrane (a) When the membrane approaches the chamber wall, valves are switched which direct the flow to the opposite side of the membrane (b)



resistance is a function of the square of the deflection [The third term of (53) on page 38 describes the strain of the neutral fiber.] and the displaced volume. The output voltage as a function of the displaced volume  $V_w$  is obtained by inserting the deflection from (303) into (451) (page 478):

$$U_m = \frac{2}{3} U_0 \frac{V_w^2}{\pi^2 R_M^8} ((1 - \nu)r_t^2 + (1 + \nu)r_r^2). \tag{484}$$

Similarly, the output voltage as a function of the displaced volume  $V_w$  is calculated for a monocrystalline silicon membrane etched into a (100) wafer with p-doped stain gauges on the surface. The strain gauges are orientated parallel and perpendicular to the edges of the square membrane. In this case, the displaced volume is described by (306) on page 170. The membrane deflection calculated with that equation is then inserted into (457) (page 281):

$$U_m = 43.5 \frac{U_0 E_M d_M V_w}{a_M^6} \times \frac{\pi_t [r_t^2(1 + 3 \nu_M) - r_r^2(3 + \nu_M)] - \pi_r [r_r^2(1 + 3 \nu_M) - r_t^2(3 + \nu_M)]}{1 - \nu_M^2}. \tag{485}$$

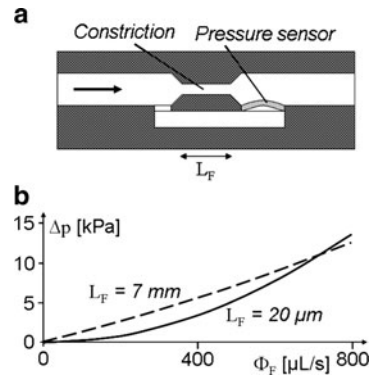
There are two ways how to drive this kind of sensor: It is possible to detect the position of the membrane with the strain gauges as a function of time or to count the number of valve switchings in a certain time. The latter method requires simpler electronics, while the former one is more accurate. Every switching of the valves results in a small flow generated by the difference in the dead volumes of the valves. Maybe the flow generated by the difference in the dead volumes is still detected and no faulty measurement occurs, but a flow sensor should not generate any flow. Therefore, a chamber with a large diameter is desirable which does not require frequent valve switching.

Another important principle of flow sensors is to measure the *pressure drop over a capillary* or constriction. Figure 220a shows a schematic drawing of this type of flow sensor. The pressure drop can be calculated with (188) on page 112 which consists of two terms. The first term is called the Torricelli equation and is a quadratic function of the volume flow  $\Phi_F$ . The second term is the Hagen Poiseuille equation and a linear function of volume flow:

$$\Delta p = -\chi \frac{\rho_F}{2} \left( \frac{1}{A_1^2} - \frac{1}{A_2^2} \right) \Phi_F^2 - 32 \frac{\eta L_F}{D_h^2} \frac{\Phi_F}{A_2}. \tag{188}$$

The viscosity  $\eta$  of the fluid is a strong function of temperature. Therefore, the Hagen Poiseuille equation yields a sensor with a linear characteristic curve sensitive to temperature changes, while the Torricelli equation results in small cross sensitivity and a nonlinear characteristic curve. This is shown in Fig. 220b by an

**Fig. 220** Measurement of the pressure difference over a constriction (a) and characteristic curves of such a sensor with a short ( $20\ \mu\text{m}$ ) and a long ( $7\ \text{mm}$ ) constriction (b)



example calculation with (188). In both cases, viscosity, density, radius of the circular flow channel, and the quantity  $\chi$  were assumed to be  $1\ \text{mPa s}$ ,  $1,000\ \text{kg/m}^3$ ,  $200\ \mu\text{m}$ , and  $0.64$ , respectively. In one case, a radius and length of the constriction of  $195\ \mu\text{m}$  and  $7\ \text{mm}$ , respectively, have been used and in the other case  $180\ \mu\text{m}$  and  $20\ \mu\text{m}$ , respectively. It is clearly seen that the longer constriction results in a more linear characteristic curve of the sensor. Thus, by designing length  $L_F$  of the constriction and cross-sectional areas of constriction  $A_2$  and flow channel  $A_1$  either a linear or a less cross-sensitive flow sensor can be obtained [77].

## Exercises

### Problem 52

In a flow sensor, there is mounted in the middle of the flow channel a  $20\text{-}\mu\text{m}$  thick carrier foil from polyimide with conductor paths from gold. In the following table, all data required for the calculations are shown.

Total length of the conductor path of the heater	$200\ \mu\text{m}$	Density of conductor paths	$19,000\ \text{kg/m}^3$
Effective distance of conductor paths to the rim of the carrier foil	$220\ \mu\text{m}$	Specific heat capacitance of conductor paths	$0.13\ \text{kJ}/(\text{kg K})$
Distance between heater and measurement wire	$200\ \mu\text{m}$	Specific heat capacitance of carrier membrane	$1\ \text{kJ}/(\text{kg K})$
Width of conductor paths	$5\ \mu\text{m}$	Heat conductivity of water	$0.6\ \text{W}/(\text{K m})$
Thickness of conductor paths	$100\ \text{nm}$	Dynamic viscosity of air at $20^\circ\text{C}$	$18.2\ \mu\text{Pa s}$
Density of carrier membrane	$1.1\ \text{kg/m}^3$	Density of water	$1,000\ \text{kg/m}^3$
Electrical resistance of each conductor path	$10\ \Omega$	Allowed temperature enhancement	$10\ \text{K}$

(continued)



Width of channel	200 $\mu\text{m}$	Heat conductivity of air	0.025 W/(K m)
Height of channel	200 $\mu\text{m}$	Density of air	1.20 kg/m <sup>3</sup>
Temperature of medical	20°C	Flow rate	2.40 $\mu\text{L}/\text{min}$
Dynamic viscosity of water at 20°C	1 mPa s		

- (a) For dosing of a medical, first the flow velocity shall be measured by a conductor path as an anemometer. The maximum temperature difference to the environment shall be limited to 10 K because otherwise the medical could be damaged. The physical properties of the medical are the same as those of water. The constant (a) was measured before to be 0.2. Please calculate the heat power transported away from the wire by the flowing medical, by heat conductivity through the fluid to the channel walls, and the electrical power required keeping the wire at its temperature.
- (b) It cannot be made sure that there are no air bubbles in the medical. Please calculate the temperature enhancement when a bubble reaches the heater. Assume that the total area of the heater is in contact to the bubble and that the bubble extends over the entire cross-section of the channel. The electrical power of the heater shall not be changed when the bubble arrives and is constant at 130  $\mu\text{W}$ . Do not take into account for your calculation the heat stored in carrier foil and heater.
- (c) How much would the heater get hotter, if there would be no heat conductivity?
- (d) Is the sensor discussed above an anemometer or a calorimeter? (The specific heat capacity of is 4,182 J/(kg K).)
- (e) What is the reaction time of the flow sensor if it is arranged in a bridge driven at a constant electrical current of 1 mA and the temperature of the heater is 10 K above the environment? Calculate the reaction time with and without the effect of the heat capacity of the carrier membrane.

### Problem 53

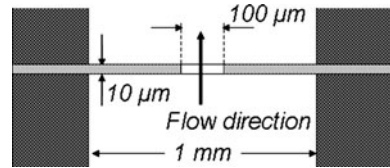
To minimize the influence of temperature on the measurement, a flow sensor as shown in Fig. E37 with a circular cross-section can be employed.

- (a) Please calculate the Hagen-Poiseuille part and the Torricelli part of the pressure difference, if water is the fluid. Do you expect a large cross sensitivity of the sensor to temperature changes?

The data needed for the calculation are shown in the following table:

Water flow to be measured	100 $\mu\text{L}/\text{min}$
Density of water	1,000 kg/m <sup>3</sup>
Viscosity of water	1 mPa s
Diameter of capillary	1 mm
Diameter of aperture	100 $\mu\text{m}$
Thickness of membrane (= Length of channel)	10 $\mu\text{m}$
$\chi$	1

**Fig. E37** Flow sensor with an aperture



- (b) Another possibility to get a flow sensor which is only a weak function of temperature is to mount instead of an aperture several wires in a flow channel and to measure the flow time of a heat pulse. What distance is required between heater and sensing wire to avoid the measurement to be disturbed by heat diffusion at the conditions at (a). What is the reaction time of the sensor?

Heat conductivity of water	0.6 W/(K m)
Specific heat capacity of water	4,182 kJ/(K kg)

### **Problem 54**

Figure 219 on page 298 shows a volume flow sensor which is independent on the kind of the fluid.

Radius of membrane	3 mm	Residual stress of membrane	100 MPa
Thickness of membrane	2 μm	Flow	10 μL/s
Young's modulus of membrane	4 GPa	Chamber depth below and above membrane	100 μm

- (a) What volume is passing through the sensor between two switchings of the valves? (Assume that the membrane is allowed just to touch bottom or ceiling of the chamber.)
- (b) The switching time of the valves needs to be much shorter than the time between two switchings of the valves, e.g., 1% of it. Otherwise, some fluid could flow uncontrolled and not detected during switching. How short must the switching time of the valves be?
- (c) How needs the design of the sensor to be changed if the valves cannot achieve the required short switching time?
- (d) What is the maximum pressure loss and pressure gain generated in the fluid because the membrane has to be deflected and tends to return into its idle state, respectively?

# Inertial Sensors

Inertial sensors are both sensors of acceleration and yaw rate (angular velocity). The first large application of *acceleration sensors* was triggering air bags in cars. Nowadays, there are a lot of other applications such as the detection of the direction in which a cellular phone is held, controllers of computer games, and detection of the position of a GPS device as long as the contact to the satellites is lost.

Nearly, all acceleration sensors employ a seismic mass fixed to a beam. The earliest design is shown in Fig. E15 on page 83 [38]. The mass is deflected when accelerated and the deflection is measured.

If the deflection is measured with strain gauges on the surface of the beam, the strain  $\varepsilon_B$  can be calculated with (98) (page 67) assuming as an approximation that the force  $F$  acting at the end of the beam in transversal direction is the mass  $m_0$  of the seismic mass multiplied by the acceleration  $a_m$ :

$$\varepsilon_B = \frac{d_B}{2} \frac{m_0(L_B - x)}{E_B I} a_m. \tag{486}$$

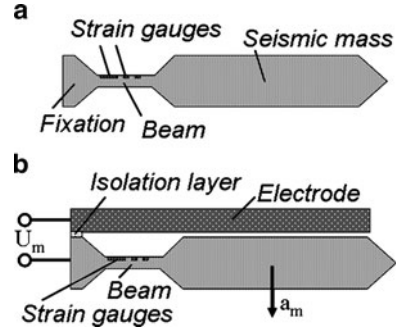
If the beam is etched into single crystalline silicon and longitudinal and transversal strain gauges are fabricated by p-doping (cf. Fig. 221a), the resistance change can be calculated with (22) on page 23:

$$\Delta R_{el} = R_{el}(\pi_l \sigma_l + \pi_b \sigma_b + \alpha_T \Delta T). \tag{22}$$

To calculate the resistance change, besides the piezoresistive coefficients in longitudinal  $\pi_l$  and transversal  $\pi_b$  direction of the strain gauges, respectively, the stresses  $\sigma_l$  and  $\sigma_b$  of the beam at the position of the strain gauges need to be known. Since the beam is fixed only in longitudinal direction, there is only longitudinal stress which can be calculated by Hooke's law. As a result, the resistance changes of longitudinal  $\Delta R_{el,l}$  and transversal  $\Delta R_{el,t}$  strain gauges on the beam are:

$$\frac{\Delta R_{el,l}}{R_{el}} = \pi_l E_B \varepsilon_B + \alpha_T \Delta T = \pi_l \frac{d_B}{2} \frac{m_0(L_B - x_l)}{I} a_m + \alpha_T \Delta T, \tag{487}$$

**Fig. 221** Schematic cross-sectional view of an acceleration sensor with (a) moving and (b) balanced seismic mass



$$\frac{\Delta R_{el,t}}{R_{el}} = \pi_t E_B \varepsilon_B + \alpha_T \Delta T = \pi_t \frac{d_B}{2} \frac{m_0(L_B - x_t)}{I} a_m + \alpha_T \Delta T. \quad (488)$$

In the above equations,  $E_B$ ,  $d_B$ ,  $L_B$ ,  $I$ , and  $x_l$ ,  $x_t$ , are Young's modulus, thickness, length, area momentum of inertia of the beam, and the effective distances to the fixation point of the beam of longitudinal and transversal strain gauges, respectively. The output signal  $U_m$  of a half bridge supplied with voltage  $U_0$  is found by inserting these equations into (447) (page 276):

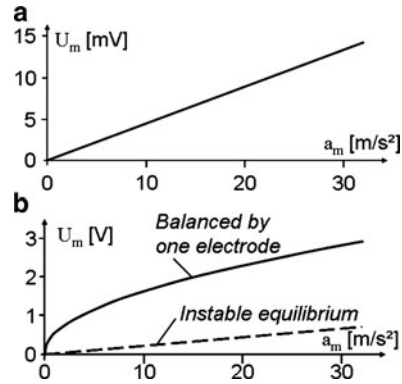
$$U_m = \frac{U_0}{4 R_{el}} (\Delta R_2 - \Delta R_1) = \frac{U_0 d_B m_0}{8 I} [(L_B - x_l)\pi_l - (L_B - x_t)\pi_t] a_m. \quad (489)$$

The piezoresistive coefficients need to be chosen according to the orientation of the beam relative to the crystal lattice of the silicon (cf. Fig. 17 on page 25).

Equation (489) is the characteristic curve of an acceleration sensor consisting of a seismic mass at the free end of a beam clamped at one end when the deflection of the mass is measured by strain gauges from silicon which are arranged longitudinal and transversal on the surface of the beam. This characteristic curve is a linear function of the acceleration. In Fig. 222a, there is shown a characteristic curve calculated with (489) with support voltage  $U_0$ , beam length  $L_B$ , width  $b_B$ , thickness  $d_B$ , seismic mass  $m_0$ , effective position of longitudinal  $x_l$  and transversal  $x_t$  strain gauges from p-silicon in  $\langle 110 \rangle$ -direction, and piezoresistive coefficients of longitudinal  $\pi_l$  and transversal  $\pi_t$  straining of 3 V, 500  $\mu\text{m}$ , 50  $\mu\text{m}$ , 20  $\mu\text{m}$ , 4.7 mg, 100  $\mu\text{m}$ , 300  $\mu\text{m}$ , 0.72/GPa, and  $-0.66/\text{GPa}$ , respectively.

In principle, it is possible to measure the deflection of a beam also with strain gauges from metal on the surface of the beam. The disadvantage of this approach is that the beam needs to be small in order to be sensitive and economic. As a consequence, the metal conductor paths cannot be very long, and, therefore, their resistance is small. Heating of the beam must be avoided, because it changes both Young's modulus and electrical resistance. Thus, the supply voltage  $U_0$  of a bridge needs to be low and only very small output signals can be achieved.

**Fig. 222** Characteristic curves of acceleration sensors with (a) moving and (b) balanced seismic mass



Acceleration sensors should be sensitive and show a large resonance frequency to allow the measurement of quick acceleration changes. The resonance frequency of a beam clamped at one end is large, if its thickness is large, its length is short, and the mass at the end of the beam is small [cf. (165) on page 98]. Unfortunately, all this results in small sensitivity according to (489).

Both high sensitivity and quick reaction time can be obtained with *balanced acceleration sensors*. In this kind of sensors, the acceleration force is balanced by some force such that the seismic mass is not moving. If the mass always (nearly) stays in its idle position, the properties of the beam such as Young’s modulus, geometry, and resonance frequency do not influence the performance of the sensor. This is the same principle as described for pressure sensors on page 282f. As shown in Fig. 221b, the seismic mass can be held, e.g., by an electrostatic force. The position of the mass can be measured by strain gauges on the beam or by a small alternating voltage applied to the electrodes in series with another capacitor or resistance. The applied alternating voltage should not change the beam deflection much. Thus, it may be an advantage to employ small electrodes to which a larger alternating voltage can be applied without significantly moving the beam.

Inertial force and electrostatic force are equal, if the beam is in its idle position. Thus, the voltage  $U_m$  required to balance acceleration and electrostatic force can be calculated from (222) (page 131):

$$\frac{1}{2} \epsilon_0 \epsilon_r \frac{A_C}{d_C^2} U_m^2 = a_m m_0 \Rightarrow U_m = d_C \sqrt{\frac{2 m_0 a_m}{\epsilon_0 \epsilon_r A_C}}. \quad (490)$$

The above equation shows that the voltage required holding the beam in its idle position is only a function of the area  $A_C$  of the electrodes, its distance  $d_C$ , and (approximately) the mass  $m_0$  at the end of the beam. It is no function of the properties of the beam (besides the fact that the mass of the beam also is accelerated). As a consequence, fabrication tolerances can be larger. The reaction time of the sensor is a function of the sensitivity of the position measurement, flexibility of

the beam, and how fast the electronic is which keeps the beam in place. As a consequence, the beam should be designed thin and long and with a comparatively large mass at its end.

The characteristic curve of an acceleration sensor with a seismic mass held by the electrostatic force in the idle position of the beam as calculated with (490) is shown in Fig. 222b compared with the measurement of the beam displacement with strain gauges from monocrystalline silicon according to (489). Area of and distance between the electrodes were assumed to be  $2 \times 2 \text{ mm}^2$  and  $1 \text{ }\mu\text{m}$ , respectively. All other parameters were the same as for Fig. 222a.

The characteristic curve described by (490) is not linear and only accelerations in one direction can be detected. These disadvantages can be avoided by the design shown in Fig. 223a or b. When there is no acceleration, the seismic mass is put to ground potential and it is attracted both to the upper and the lower electrode by the same electrostatic force. Therefore, it is in an instable equilibrium. When acceleration occurs, the voltage  $\Delta U$  at the seismic mass is adapted such that the mass stays in its idle position.

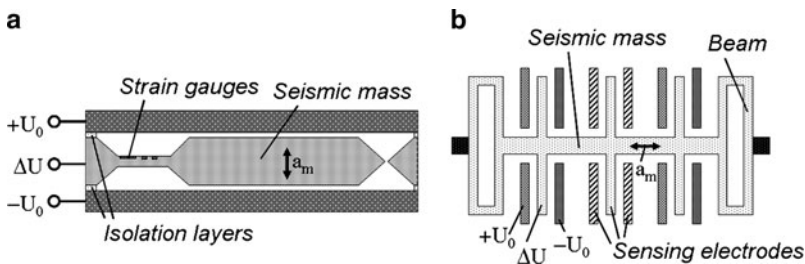
The forces acting at the seismic mass are the acceleration and the electrostatic forces pulling it up and down as calculated with (225) (page 133):

$$F_{C,n} = \frac{1}{2} \varepsilon_0 \varepsilon_r A_C \left( \frac{(U_0 + \Delta U)^2}{d_C^2} - \frac{(U_0 - \Delta U)^2}{d_C^2} \right) = 2 \frac{\varepsilon_0 \varepsilon_r A_C}{d_C^2} U_0 \Delta U = a_m m_0$$

$$\Rightarrow \Delta U = \frac{m_0 d_C^2}{2 \varepsilon_0 \varepsilon_r A_C U_0} a_m. \quad (491)$$

The above equation is the characteristic curve of sensors as shown in Fig. 223a or b. It is shown in Fig. 222b with the same parameters as used for the calculation of the seismic mass balanced by one electrode according to (490).

When an acceleration sensor is designed, as for the design of any sensor cross sensitivities need to be considered. In general, the largest cross sensitivity is due to temperature changes. If the seismic mass is fixed at the free end of only one beam, temperature and other effects such as bending or straining of the substrate do not affect the characteristic curve. However, Young's modulus of the beam is a

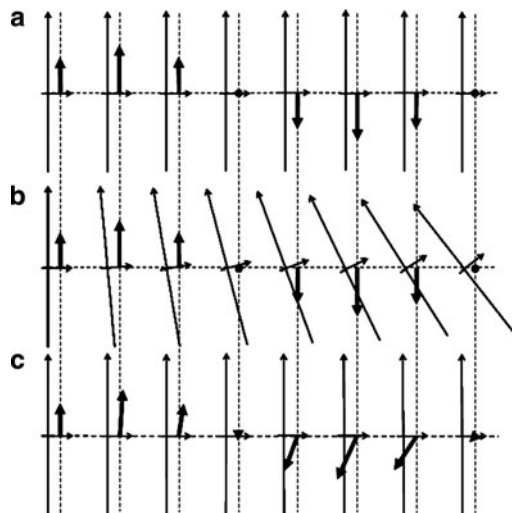


**Fig. 223** Balanced acceleration sensors. (a) Cross-section of a sensor etched into silicon, (b) top view of a sensor fabricated by surface micromachining

function of temperature. This will have no effect on balanced acceleration sensors because the properties of the beam do not change their characteristic curves.

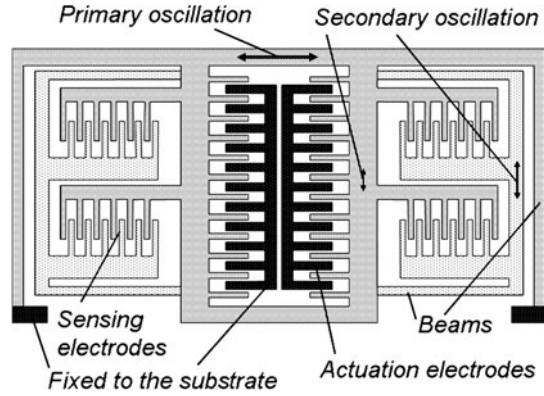
Acceleration sensors may also be sensitive to accelerations in directions other than the desired one or to angular accelerations. In Fig. 223a, the beam is in the line of the center of mass of the seismic mass. This way bending of the beam due to acceleration longitudinal to the beam is avoided. The influence of an acceleration perpendicular to the plane of the cross-section shown in Fig. 223a may be reduced by two parallel beams or a beam which is much wider than thick. Strain gauges may also be arranged such that their combination to a bridge compensates the effect of beam bending in a direction in which the sensor shall be insensitive (cf. Problem 13 on page 82). If a system of acceleration sensors measuring in all three directions is constructed, it may also be a strategy to compensate the signal from the sensor of one direction with the signals of the other directions.

*Angular velocity sensors*, also called *yaw rate sensors*, are employed mainly in automobiles for electronic stability control (ESC) and other applications in cars. The measurement of angular velocity is based on the Coriolis acceleration which appears to act on a moving body in a rotating system. The Coriolis acceleration is a fictitious acceleration because it is just the interpretation of a rotating observer who tries to understand why a mass is deflected although there appears to be no reason for that if the own rotation is not taken into account. In Fig. 224a, there is shown how a mass is oscillating back and forth (e.g., at the end of a spring) in an inertial system. Figure 224b shows the same motion but in addition there is shown a rotating system, and in Fig. 224c the oscillation is shown from the point of view of an observer rotating together with the system. In the rotating system, the mass appears to fulfill an additional secondary oscillation perpendicular to the primary one.



**Fig. 224** Oscillator (*arrow*) as a function of time in a resting (**a**) with a rotating (**b**) and observed in the rotating (**c**) coordinate system

**Fig. 225** Principle of an angular velocity sensor



This principle is employed for the measurement of angular velocity. There are a lot of ways how the Coriolis acceleration is measured. Here just an example is shown which explains the principle. In Fig. 225, there is shown the schematic top view of a yaw rate sensor. Only the dark parts are fixed to the substrate and the whole rest is movable and supported by the beams which act as springs. An alternating voltage applied to the actuator electrodes generates a primary oscillation of the entire structure. When the oscillating structure is rotating, a secondary oscillation is generated which is measured with the sensing electrodes.

The position  $x$  of the body moved by the primary oscillation as a function of time  $t$  is described by a sine function:

$$x(t) = A_0 \sin(\omega t) \Rightarrow \quad (492)$$

$$v(t) = \frac{\partial x}{\partial t} = A_0 \omega \cos(\omega t). \quad (493)$$

In the above equation,  $A_0$  is the amplitude of the primary oscillation and  $v$  is the velocity of the body due to the primary oscillation.

The Coriolis acceleration  $a_c$  in general is calculated from the cross product of the vectors of the velocity  $v$  of the moving mass and the angular velocity  $\Omega$  of the rotating system by the following equation:

$$\mathbf{a}_c = 2 \mathbf{v} \times \boldsymbol{\Omega} = 2 v \Omega \sin(\alpha). \quad (494)$$

In the above equation,  $\alpha$  is the angle between the vectors of the velocity and the angular velocity of the rotation. The sensor shown in Fig. 225 shall measure the angular velocity around an axis perpendicular to the plane shown. Therefore, the sine of  $\alpha$  is zero and inserting (493) into (494) yields:

$$a_c = 2 A_0 \omega \Omega \cos(\omega t). \quad (495)$$



Obviously, a large Coriolis acceleration can be achieved when amplitude and frequency of the primary oscillation are large. Thus, it is an advantage to drive the device at its resonance frequency. Detection is eased when the beams supporting the sensing electrodes are also oscillating at their resonance.

# References

1. Finne RM, Klein DL (1967) A water-amine-complexing agent system for etching silicon. *J Electrochem Soc* 114:965–970
2. Nathanson HC, Newell WE, Wickstrom RA, Davis JR Jr (1967) The resonant gate transistor. *IEEE Trans Electron Devices* ED-14:117–133
3. Sze SM (1985) *Semiconductor devices, physics and technology*. Wiley, New York. ISBN 0-471-83704-0
4. Elwenspoek M, Jansen H (1998) *Silicon micromachining*. Cambridge University Press, Cambridge. ISBN 0521 60767 1
5. Menz W, Mohr J, Paul O (2001) *Microsystem technology*. Wiley-VCH Verlag GmbH, Weinheim. ISBN 3-527-29634-4
6. Madou MJ (2002) *Fundamentals of microfabrication*. CRC Press, West Palm Beach, FL. ISBN 0-8493-0826-7
7. Tsai C-L, Henning AK (1997) Surface micromachined turbines. In: *Proceedings of the international conference on solid-state sensors and actuators, Transducers'97, 1997*, pp 829–832
8. Haji-Babaei J, Kwok CY, Huang RS (1997) Integrable active microvalve with surface micro-machined curled-up actuator. In: *Proceedings of the international conference on solid-state sensors and actuators, Transducers'97, Chicago, USA, 16–19 July 1997*, pp 833–836
9. Stoney GG (1909) The tension of metallic films deposited by electrolysis. *Proc R Soc Lond A* 82:172–175
10. Hunt RA, Gale B (1972) A model of deadherence due to stresses in an elastic film. *J Phys D* 5:359–372
11. Yanof AW, Resnick DJ, Jankoski CA, Johnson WA (1986) X-ray mask distortion: process and pattern dependence. *SPIE* 632:118–132
12. Flinn PA (1989) Principles and applications of wafer curvature techniques for stress measurements in thin films. In: Bravman JC, Nix WD, Barnett DM, Smith DA (eds) *Thin films: stresses and mechanical properties*, (Mater. Res. Soc. Symp. Proc. Volume **130**, Pittsburgh, PA 1989), pp 41–51
13. Wagner B, Quenzer HJ, Hoerschelmann S, Lisek T, Juers M (1996) Bistable microvalves for implantable drug delivery systems. In: *Proceedings of the 5th international conference on new actuators, ACTUATOR'96, Bremen, 1996*, pp 33–36
14. Zook JD, Burns DW, Guckel H, Sniegowski JJ, Engelstad RL, Feng Z (1992) Characteristics of polysilicon resonant microbeams. *Sens Actuators A* 35:51–59
15. van Lintel HTG, van de Pol FCM (1988) A Piezoelectric micropump based on micromachining of silicon. *Sens Actuators A* 15:153–167
16. Thomas ME, Hartnett MP, McKay JG (1988) The use of surface profilometers for the measurement of wafer curvature. *J Vacuum Sci Technol A* 6:2570–2571
17. Schmidt OG (2001) Max-Planck-Gesellschaft, Presseinformationen

18. Kanda Y (1982) A graphical representation of the piezoresistance coefficients in silicon. *IEEE Trans Electron Devices* ED29:64–70
19. Timoshenko S, Woinowsky-Krieger S (1959) *Theory of plates and shells*, 2nd edn. McGraw-Hill, New York
20. Schomburg WK, Rummeler Z, Shao P, Wulff K, Xie L (2004) The design of metal strain gauges on diaphragms. *J Micromech Microeng* 14:1101–1108
21. Bronstein IN, Semendjajew KA (1976) *Taschenbuch der Mathematik*, 17 Auflage. ISBN 3 87144 0167
22. Beams JW (1959) Mechanical properties of thin films of gold and silver. In: Neugebauer CA, Newkirk JB, Vermilyea DA (eds) *Structures and properties of thin films*. Wiley, New York, pp 183–192
23. Pan JY, Lin P, Masceh F, Senturia SD (1990) Verification of FEM analysis of load deflection methods for measuring mechanical properties of thin films. *Proceedings of the Hilton Head Island, SC, USA, 4–7 June 1990*, pp 951–956
24. Ritz W (1909) Über eine neue Methode zur Lösung gewisser Variationsprobleme der Mathematischen Physik. *J reine angewandte Math* 135:1–61
25. Goll C, Bacher W, Büstgens B, Maas D, Menz W, Schomburg WK (1996) Microvalves with bistable buckled polymer diaphragms. *J Micromech Microeng* 6:77–79
26. Schomburg WK, Ahrens R, Bacher W, Engemann S, Krehl P, Martin, J (1997) Long-term performance analysis of thermo-pneumatic micropump actuators. In: *Proceedings of the 9th international conference on solid-state sensors and actuators, transducers'97, Chicago, USA, 16–19 June 1997*, pp 365–368
27. Di Giovanni M (1982) *Flat and corrugated diaphragm design handbook*. Marcel Dekker, New York. ISBN 0-8247-1281-1
28. Haringx JA (1957) Design of corrugated diaphragms. *ASME Trans* 79:55–64
29. Wortman JJ, Evans RA (1965) Young's modulus, shear modulus, and Poisson's ratio in silicon and germanium. *J Appl Phys* 36:153–156
30. Schomburg WK, Goll C (1998) Design optimization of bistable microdiaphragm valves. *Sens Actuators A* 64:259–264
31. Kawle A (2009) Equation for the lateral deflection of a folded beam derived by FEM calculations. Mini thesis at RWTH Aachen University, KEMikro
32. Blevins RD (1995) *Formulas for natural frequency and mode shape*. R.E. Krieger, Malabar, FL, ISBN/ISSN 0-89464-894-2
33. Beitz W, Grote K-H (2001) *Dubbel – Taschenbuch für den Maschinenbau*. Springer, Heidelberg. ISBN 3-540-67777-1
34. Czichos H (ed) (1991) *Hütte – Die Grundlagen der Ingenieurwissenschaften*. Springer, Heidelberg. ISBN 3-540-19077-5 and 0-387-190077-5
35. Figure obtained from [http://en.wikipedia.org/wiki/Golden\\_Gate\\_Bridge](http://en.wikipedia.org/wiki/Golden_Gate_Bridge), 21 Dec 2010
36. Strohrmann M, Mohr J, Schulz J (1995) *Wissenschaftliche Berichte FZKA* 5561
37. Chen H, Bao M, Zhu H, Shen S (1997) A piezoresistive accelerometer with a novel vertical beam structure. *Transducers* 97:1201–1204
38. Roylance LM, Angell JB (1979) A batch-fabricated silicon accelerometer. *IEEE Trans Electron Devices* 26:1911–1917
39. Schomburg WK, Vitt M, Bacher W, Börner MW, Menz W (1995) Measurements of physical parameters with ultrasound and microdiaphragms. In: *Proceedings of micro electro mechanical systems'95*, 29 Jan–2 Feb 1995, Amsterdam, pp 139–144
40. Rayleigh L (1945) *Theory of sound (1877)*, vol 1, Newth edn. Dover Publications, New York
41. den Hartog JP (1956) *Mechanical vibrations*, 4th edn. McGraw-Hill, New York
42. Lamb H (1920) On the vibrations of an elastic plate in contact with water. *Proc R Soc Lond A* 98:205–216
43. Shaker FJ (1975) Effect of axial load on mode shapes and frequencies of beams. Lewis Research Center Report NASA-TN-8109

44. Benecke W, Csepregi L, Heuberger A, Kühl K, Seidel H (1985) A frequency-selective, piezoresistive silicon vibration sensor. *Transducers'85*, Philadelphia, June 1985, p 105
45. Trah H-P, Baumann H, Döring C, Goebel H, Grauer T, Mettner M (1993) Micromachined valve with hydraulically actuated membrane subsequent to a thermoelectrically controlled bimorph cantilever. *Sens Actuators A* 39:169–176
46. Vollmer J, Hein H (1994) Miniaturisierte fluidische Strahlelemente in LIGA-Technik mit integrierter elektrischer Steuerung. *Berichte Kernforschungszentrum Karlsruhe* **5375**
47. Menz W, Mohr J, Paul O (2005) *Mikrosystemtechnik für Ingenieure*, p 406
48. Lenk A, Pfeifer G, Werthschützky R (2001) *Elektromechanische Systeme*. Springer, Heidelberg. ISBN 3-540-6941-3
49. Pfeifer G (1982) Piezoelektrische lineare Stellantriebe. *Wissenschaftliche Schriftenreihe der Technischen Hochschule Karl-Marx-Stadt*, 6. ISSN 0323–6374
50. Destruel P, Rojas FS, Tougne D (1984) Hoang-the-giam. *J Appl Phys* 56:3298–3303
51. Franz J, Baumann H, Trah H-P (1996) A monolithic silicon micro-liquid handling system with bimetallic valve and calorimetric flow sensor. In: *Proceedings of the 5th international conference on new actuators, ACTUATOR'96*, Bremen, 1996, pp 29–32
52. Lisec T, Hoerschelmann S, Quenzer HJ, Wagner B, Benecke W (1994) A fast switching silicon valve for pneumatic control systems. In: *Proceedings of the 4th international conference on new actuator, Actuator'94*, 1994, pp 30–33
53. Saleh BEA, Teich MC (1991) *Fundamentals of photonics*. Wiley, New York
54. Clusser EL (1984) *Diffusion – mass transfer in fluid systems*. Cambridge University Press, Cambridge
55. Stroock AD, Dertinger SKW, Ajdari A, Mezic I, Stone HA, Whitesides GM (2002) Chaotic mixer for microchannels. *Science* 295:647–51
56. Ehrhard P, Meisel I, Farangis Zadeh H (2004) FEM-Simulationen und Validierungs-Experimente zu elektrisch erregten Mikromischern. 5. Statuskolloquium des Programms Mikrosystemtechnik, 10–11 Feb 2004, pp 165–169
57. Khuntontong P, Blaser T, Maas D, Schomburg WK (2008) Fabrication of a polymer micro mixer by ultrasonic hot embossing. In: *Proceedings of the 19th microMechanics Europe workshop, MME 2008*, 28–30 Sept 2008, pp 259–262. ISBN 978-3-00-025529-8
58. Henning AK (2003) Improved gas flow model for microvalves. In: *Transducers'03*, pp 1550–1553
59. Rogge T, Rummler Z, Schomburg WK (2004) Polymer micro valve with a hydraulic piezo-drive fabricated by the AMANDA process. *Sensors and Actuators A* 110:206–212
60. Jerman H (1991) Electrically-activated, normally-closed diaphragm valves. In: *Transducers'91*, pp 1045–1048
61. Shikida M, Sato K (1994) Characterization of an electrostatically-driven gas valve under high pressure conditions. In: *Proceedings of the IEEE MEMS*, pp 235–240
62. Wagner B, Quenzer HJ, Hoerschelmann S, Lisec T, Juerss M (1996) Bistable microvalves for implantable drug delivery systems. In: *Proceedings of the 5th international conference on new actuators, ACTUATOR'96*, 1996, pp 33–36
63. Huff MA, Gilbert JR, Schmidt MA (1993) Flow characteristics of a pressure-balanced microvalve. In: *Proceedings of the transducers*, 1993, pp 98–101
64. Zengerle R, Ulrich J, Kluge S, Richter M, Richter A (1995) A bidirectional silicon micro-pump. *Sensors and Actuators A* 50:81–88
65. Cabuz C, Herb WR, Cabuz EI, Lu ST (2001) The dual diaphragm pump. In: *Proceedings of MEMS*, 2001, pp 519–522
66. Effenhauser CS, Hartig H, Krämer P (2002) An evaporation-based disposable micropump concept for continuous monitoring applications. *Biomed Microdevices* 4(1):27–32
67. Richter A, Plettner A, Hofmann KA, Sandmaier H (1991) A micromachined electrohydrodynamic (EHD) pump. *Sens Actuators A* 29:159–168

68. Truckenmüller R, Ahrens R, Cheng Y, Fischer G, Saile V (2006) An ultrasonic welding based process for building up a new class of inert fluidic microsensors and -actuators from polymers. *Sens Actuators A* 132:385–392
69. Figures obtained from [www.debiotech.com/products/msys/mip.html](http://www.debiotech.com/products/msys/mip.html), 21 Dec 2010
70. Majumder S, McGruer NF, Zavracky PM, Adams GG, Morrison RH, Krim J (1997) Measurement and modeling on surface micromachined, electrostatically actuated microswitches. In: *Proceedings of Transducers'97*, vol. 2, 1997, pp 1145–1148
71. Landau LD, Lifshitz EM (1981) *Hydrodynamik*, 4th edn. Akademie Verlag, Berlin
72. Elwenspoek M, Wiegerink R (2001) *Mechanical microsensors*. Springer, Heidelberg, 3-540-67582-5
73. Lomas CG (1986) *Fundamentals of hot wire anemometry*. Cambridge University Press, Cambridge. ISBN 0 521 30340 0
74. Gogoi BP, Mastrangelo CH (1999) A low-voltage force-balanced pressure sensor with hermetically sealed servomechanism. In: *MEMS'99, USA, Orlando, FL, 1999*, pp 493–498
75. Ashauer M, Glosch H, Hedrich F, Hey N, Sandmaier H, Lang W (1999) Thermal flow sensor for liquids and gases based on combinations of two principles. *Sens Actuators* 73:7–13
76. Shao P, Rummeler Z, Schomburg WK (2003) Dosing system for the nano liter range, fabricated with the AMANDA process. *J Micromech Microeng* 13:85–90
77. Richter M, Wackerle M, Woias P, Hillerich B (1999) A novel flow sensor with high time resolution based on differential pressure principle. In: *MEMS'99, USA, Orlando, FL 1999*, pp 118–123
78. Zappe H (2010) *Fundamentals of Micro-Optics*, Cambridge University Press, Cambridge. ISBN 978-0521895422

#### Characteristic curves of actuators

Actuator	Figure	Page	Actuator	Figure	Page
Bimaterial	118	162	Thermal	116	160
Capacitive/electrostatic	98	134	Thermal, bimaterial	118	162
Electroosmotic pump	91	125			
Phase transition	126	171	Thermomechanic (heated beam)	123b	166
Piezo bimorph	105	145	Thermo-pneumatic	125a	169
Piezo general	111	150	Valve, active	157	208
Piezo, single	101	140	Valve, passive	171	221
Pump	182	238			

#### Characteristic curves of sensors

Sensor	Figure	Page
Piezo	102	142
Flow (anemometer)	216	292
Flow (calorimeter)	217c	295
Flow (flow time)	218	297
Flow (pressure drop over constriction)	220b	300
Pressure (metal strain gauges)	208	278
Pressure (silicon strain gauges)	210	282
Pressure (balanced)	212	283
Pressure (frequency analog)	214	286
Acceleration	222	305

# Index

## A

- Acceleration sensor
  - balanced, 305
  - general, 303
- Actuator
  - bimaterial, 161
  - capacitive, 131
  - comb drive, 132
  - di-electrophoresis, 127
  - electroosmosis, 122
  - electrostatic, 131
  - heated beam, 164
  - phase transition, 171
  - piezo, 137
  - piezo bimorph, 144
  - piezo bimorph with carrier, 148
  - piezo, circular plate, 149
  - piezo on carrier, 146
  - pump, 229
  - thermal, 159
  - thermo-pneumatic, 167
  - valves, 203
- Additive mass, 92
- Adhesion of thin films, 10
- Analogies of physical domains, 253–259
- Angular velocity sensor, 307

## B

- Beam
  - bi-stable, 75
  - clamped at both ends, 71
  - compressive stress, 75
  - critical stress, 74
  - deflected in parallel, 76
  - folded, 77
  - frequency parameters, 96, 98
  - general, 65
  - loaded with own weight, 69

- torsional deflection, 78
- vibrations, 94
- Bending, 5
- Bending of thin films, 9–15
- Bernoulli equation, 109
- Bimaterial actuator, 161
- Bimorph
  - piezo, 144
  - thermal, 161
- Bi-stable
  - beam, 75
  - membrane, 39
- Blaze, 179
- Brownian motion, 192
- Bubble, 118
- Bubble in pump chamber, 241
- Burst pressure, 44

## C

- Cabrera's equation, 33
- Capacitive force, 131–135
- Capillary
  - electroosmosis, 122
  - flow resistance, 109
  - force, 114
  - friction, 107
  - gas bubble, 118
  - general, 107
  - hydraulic diameter, 107
  - squeeze film effect, 119
  - stokes friction, 122
  - stop groove, 117
- Capillary force, 114
- Coanda effect, 113
- Comb drive, 132
- Compressive stress
  - beams, 75
  - in membrane, 39

Conductor paths, 21–26  
 Corrugated membrane, 48  
 Critical distance, 295  
 Critical flow velocity, 295  
 Critical stress  
   beam, 71, 74, 75  
   membrane, 41, 42, 92  
 Crystalline structure, 13  
 Crystal orientation, 23

**D**

Dead volume, 217, 221  
 Debye-Hückel length, 122  
 Deflection  
   beams, 65  
   beams, longitudinal, 70  
   circular thin membrane, 30  
   membranes, 50  
   thick membranes, 34  
   thin membranes, 33  
 Di-electrophoresis, 127  
 Diffraction grating, 177  
 Diffusion, 191–200  
 Dosing, 251  
 Double-slit, 177

**E**

Elastic deformations, 5–7  
 Electrical spring constant, 267  
 Electro-mechanical filter, 263  
 Electro-mechanical switch, 261  
 Electro-optical effect, 187  
 Electroosmosis, 122  
 Electroosmotic pump, 125  
 Electrophoresis, 126  
 Electrostatic force, 131–135  
 Evanescent field, 186

**F**

Fabry-Perot interferometer, 182  
 Flexible membrane, 30  
 Flow sensor  
   anemometer, 289  
   calorimeter, 295  
   critical distance, 295  
   critical flow velocity, 295  
   flow time measurement, 296  
   general, 289  
   King's law, 290  
   pressure drop, 299  
   volume displacement, 298

**G**

Gauge factor, 22

**H**

Hagen Poiseuille equation, 107  
 Heated beam, 164  
 Heat exchanger, 197  
 Hückel equation, 126  
 Hydraulic diameter, 107

**I**

Ink jet printer, 251  
 Interference, 178

**K**

Kerr effect, 187  
 King's law, 290

**L**

Lotus effect, 119

**M**

Mach-Zehnder interferometer, 187  
 Membrane  
   bi-stable, 39  
   with a boss, 47  
   with compressive stress, 39  
   corrugated, 48  
   flexible, 30  
   general, 29  
   thick, 29  
   thin, 29  
   vibrations, 85  
 Micro membrane sensors, 87  
 Micro mixers, 192  
 Micro optics, 177–189  
 Miller indices, 23  
 Mono-crystalline silicon  
   general, 23  
   piezo-resistive coefficient, 23, 25  
   Poisson's ratio, 50  
   shear modulus, 79  
   strain gauges, 61  
   Young's modulus, 50

**N**

Numerical aperture, 184

**O**

Optical fiber, 184  
 Optical wave guide, 184

**P**

Phase transition actuator, 171  
 Piezo  
   bimorph, 144  
   bimorph with carrier, 148



circular plate, 149  
 ferroelectric, 137  
 general, 137  
 on inactive carrier, 146  
 pyroelectric effect, 153  
 sensor, 141  
 Piezoelectric effect, 137–156  
 Piezo-resistive coefficient, 23  
 Pockels effect, 187  
 Polymerase chain reaction, 199  
 Pressure sensor  
   balanced, 282  
   capacitive, 282  
   frequency analog, 284  
   general, 275  
   metal strain gauges, 276  
   silicon strain gauges, 280  
 Pull in voltage, 262  
 Pump  
   with active valves, 243  
   aperiodic, 245  
   electroosmotic, 125  
   fluidic power, 239  
   gas bubble, 241  
   general, 229  
   in series, 241  
   valveless, 245

**R**

Rayleigh method, 89  
 Recrystallization, 14  
 Refraction, 183  
 Relay, 261  
 Residual stress  
   beam, 73  
   beam vibrations, 100  
   compressive, 39  
   general, 5  
   membrane, 31, 38, 91  
   thin films, 9  
 Response time, 218  
 Reynolds' number, 191

Ritz method, 34  
 Rowland circle, 181

**S**

Scaling laws, 3–4  
 Snell's law, 183  
 Squeeze film effect, 119  
 Standing wave, 182  
 Step grid, 179  
 Stokes friction, 122  
 Stop groove, 117  
 Strain gauges  
   on beams, 68  
   general, 21  
   on membranes, 53  
   from metal, 54  
   from mono-crystalline silicon, 61  
 Stress due to straining, 5  
 Stress gradient, 10  
 Super hydrophobic surface, 119

**T**

Thermal actuators, 159–173  
 Thermo-pneumatic actuator, 167  
 Thick membrane, 29  
 Thin films, 9–15  
 Thin membrane, 29  
 Total reflection, 183

**V**

Valve  
   active, 203  
   passive, 219  
 Vibrations  
   additive mass, 92  
   beams, 94  
   damping, 92  
   emission of sound, 93  
   general, 85  
   membranes, 85

**Y**

Yaw rate sensor, 307