1 Introduction

Worldwide, sensor development is a fast growing discipline. Today's sensor market offers thousands of sensor types, for almost every measurable quantity, for a broad area of applications, and with a wide diversity in quality. Many research groups are active in the sensor field, exploring new technologies, investigating new principles and structures, aiming at reduced size and price, at the same or even better performance.

System engineers have to select the proper sensors for their design, from an overwhelming volume of sensor devices and associated equipment. A well motivated choice requires thorough knowledge of what is available on the market, and a good insight in current sensor research to be able to anticipate forthcoming sensor solutions.

This introductory chapter gives a general view on sensors — their functionality, the nomenclature and global properties — as a prelude to a more in-depth discussion about sensor performance and operation principles.

1.1 Sensors in Mechatronics

1.1.1 Definitions

A transducer is an essential part of any information processing system that operates in more than one physical domain. These domains are characterized by the type of quantity that provides the carrier of the relevant information. Examples are the optical, electrical, magnetic, thermal and mechanical domains. A transducer is that part of a measurement system that converts information about a measurand from one domain to another, ideally without information loss.

A transducer has at least one input and one output. In measuring instruments, where information processing is performed by electrical signals, either the output or the input is of electrical nature (voltage, current, resistance, capacitance and so on), whereas the other is a non-electrical signal (displacement, temperature, elasticity and so on). A transducer with a non-electrical *input* is an *input transducer*, intended to convert a non-electrical quantity into an electrical signal in order to measure that quantity. A transducer with a non-electrical *output* is called an *output*

Sensors for Mechatronics. DOI: 10.1016/B978-0-12-391497-2.00001-7

		Interrogating input	
		Design controlled	Environment controlled
LIP input	Design controlled	Source	Direct sensor
	Environment controlled	Modulating sensor	Multiplying devices

Figure 2.4 Unified transducer classification.

According to the 'unified transducer model' as introduced in [7], an input port can be controlled either by design (it has a fixed value) or by the environment (the measurand or some unwanted input variable). So we have four different cases (Figure 2.4). The characteristics of these four transducer types are briefly reviewed:

- Design-controlled LIP input and design-controlled interrogating input.
  All inputs are fixed. This type represents a signal or information source, for instance a standard or a signal source with a constant or predetermined output. The output is totally determined by the construction and the materials that have been chosen. Any environmental effect on the output is (ideally) excluded.
- Design-controlled LIP input and environment-controlled interrogating input.
  Since the latent information parameters are fixed by design, the output depends only on what is connected to the interrogating input. When this is the measurand, the transducer behaves as a direct sensor. Examples:
  - *Thermocouple temperature sensor*: the Seebeck coefficient is fixed by the choice of the materials.
  - *Piezoelectric accelerometer*: the sensitivity is fixed by the seismic mass and the piezoelectric properties of the crystal.
- Environment-controlled LIP input and design-controlled interrogating input.
  The measurand affects particular material properties or geometric parameters. These changes are interrogated by a fixed or well-defined signal at the interrogating input. The transducer behaves as a modulating sensor. Examples:
  - *Strain gauge bridge*: strain alters the resistance of the strain gauges; a bridge voltage converts this resistance change into an output voltage;
  - *Linear variable differential transformer (LVDT)*: a displacement of an object connected to the moving core will change the transfer ratio of the differential transformer. An AC signal on the primary coil acts as interrogating quantity.
  - *Hall sensor*: the measurand is a magnetic induction field, which acts on moving charges imposed by a fixed (or known) current applied to the interrogating input.
- Environment-controlled LIP input and environment-controlled interrogating input.
  These are multiplying transducers: the output depends on the quantities at both inputs, often in a multiplicative relation. For instance a Hall sensor could act as such, when the interrogating input is not a fixed current (by design) but a current that is related to just another measurand.

It is important to note that any practical transducer shows all four types of responses. A strain gauge (a modulating transducer) produces, when interrogated,

3 Uncertainty Aspects

No sensor is perfect. The mechatronic designer must be aware of the sensor's shortcomings in order to be able to properly evaluate measurement results and to make a correct assessment of the system performance. Specifying sensor quality in terms of accuracy only is not sufficient: a larger number of precisely defined parameters is necessary to fully characterize the sensor's behaviour. Often a designer can reduce the effects of the intrinsic sensor limitations by the application of special configurations, procedures and methods. Similar measures can also be considered when environmental influences should be eliminated. This chapter reviews the most important terms to express sensor behaviour and presents some general design methods to reduce errors due to sensor deficiencies and environmental factors.

3.1 Sensor Specification

Imperfections of a sensor are usually listed in the data sheets provided by the manufacturer. These sensor specifications inform the user about deviations from the ideal behaviour. The user must accept technical imperfections, as long as they do not exceed the specified values.

Any measuring instrument, and hence any sensor, has to be fully specified with respect to its performance. Unfortunately, many data sheets show lack of clarity and completeness. Gradually, international agreements about formal error descriptions are being established. An exhaustive description of measurement errors and error terminology can be found in [1], along with an international standard on transducer nomenclature and terminology [2]. Various international committees are working towards a uniform framework to specify sensors [3]. Finally, a special document is in preparation, containing definitions of measurement-related terms: the International Vocabulary of Basic and General Terms in Metrology (short VIM) [4].

The characteristics that describe sensor performance can be classified into four groups:

- *Static characteristics*, describing the performance with respect to very slow changes.
- *Dynamic characteristics*, specifying the sensor response to variations in time and in the measurand (the quantity that has to be measured).
- *Environmental characteristics*, relating the sensor performance after or during exposure to specified external conditions (e.g. pressure, temperature, vibration and radiation).
- *Reliability characteristics*, describing the sensor's life expectancy.

Sensors for Mechatronics. DOI: 10.1016/B978-0-12-391497-2.00003-0

Another important type of noise is *1/f noise* (one-over-f noise), a collection of noise phenomena with a spectral noise power that is proportional to f^{-n}, with $n = 1-2$.

Quantization noise is the result of quantizing an analogue signal. The rounding off results in a (continuous) deviation from the original signal. This error can be considered as a 'signal' with zero mean and a standard deviation determined by the resolution of the AD converter.

3.1.7 Response Time

The response time is associated with the speed of change in the output on a stepwise change of the measurand. The specification of the response time needs *always* be accompanied with an indication of the input step (for instance FS — full scale) and the output range for which the response time is defined, for instance 10–90%. Creep and oscillations may make the specification of the response time meaningless or at least misleading.

3.1.8 Frequency Response and Bandwidth

The sensitivity of a system depends on the frequency or rate of change of the measurand. A measure for the useful frequency range is the frequency band. The upper and lower limits of the frequency band are defined as those frequencies for which the output signal has dropped to half the nominal value, at constant input power. For voltage or current quantities the criterion is ½$\sqrt{2}$ of the nominal value. The lower limit of the frequency band may be zero; the upper limit has always a finite value. The extent of the frequency band is called the *bandwidth* of the system, expressed in Hz.

3.1.9 Operating Conditions

All specification items only apply within the operating range of the system, which should also be specified correctly. It is given by the measurement range, the required supply voltage, the environmental conditions and possibly other parameters.

Example 1

The frequency characteristics and noise behaviour of an accelerometer are important features. Table 3.1 is an excerpt from the data sheets of the QA-2000 accelerometer from Allied Signal Aerospace.

Example 2

Many humidity sensors have non-linear behaviour. Table 3.2 is an example of the specifications of a humidity sensor EMD 2000 from Phys-Chem SCIENTIFIC Corp., NY; RH stands for relative humidity.

Despite the specified limitations of sensors, a sensing system can be configured in a way that the effect of some of these limitations are eliminated or at least reduced. We will consider various possibilities of error-reducing designs in the next section.

Table 3.1 Selected Specifications of an Accelerometer

Environmental:	Temperature	−55°C to +95°C
	Shock	250 g, half-sine, 6 ms
	Vibration	MIL-E-5400 curve IV(A)
Frequency response:	0–10 Hz	0.01 dB
	10–300 Hz	0.45 dB
	Above 300 Hz	$<$5 dB peaking
	Natural frequency	$>$800 Hz
Noise:	0–10 Hz	10 nA rms
	10–500 Hz	100 nA rms
	500 Hz to 10 kHz	2 μA rms

Table 3.2 Selected Specifications of a Humidity Sensor

Operating RH range	10–98% RH
Operating temperature range	−10°C to 75°C
Response time	10 s for a step change from 11% RH to 93% to reach 90% or better of equilibrium value
Hysteresis	±0.3% RH at 25°C

Note: Step response and hysteresis curves are included in the specification sheets.

3.2 Sensor Error Reduction Techniques

Any sensor system has imperfections, introducing measurement errors. These errors either originate from the system itself (for instance system noise, quantization and drift) or are due to environmental influences such as thermal, electromagnetic and mechanical interference. Sensor manufacturers try to minimize such intrinsic errors through proper design of the sensor layout and encapsulation; the remaining imperfections should be given in the data sheets of the sensor. The user (for instance the mechatronic designer) should minimize additional errors which could arise from improper mounting and faulty electronic interfacing. In this section we present some general concepts to minimize or to reduce the effect of the intrinsic errors when applying sensors.

Usually, a sensor is designed to be sensitive to just one specific quantity, thereby minimizing the sensitivity to all other quantities, despite the unavoidable presence of many physical effects. The result is a sensor that is sensitive not only to the quantity to be measured but also in a greater or lesser degree to other quantities; this is called the cross-sensitivity of the device. Temperature is feared most of all, illustrated by the saying that 'every sensor is a temperature sensor'.

Besides cross-sensitivities, sensors may suffer from many other imperfections. They influence the transfer of the measurement signal and give rise to unwanted output signals. Figure 3.1 shows a simplified model of a sensor system, with an

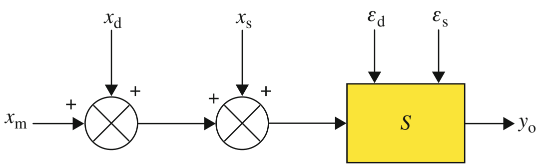

Figure 3.1 Simplified sensor model including error signals.

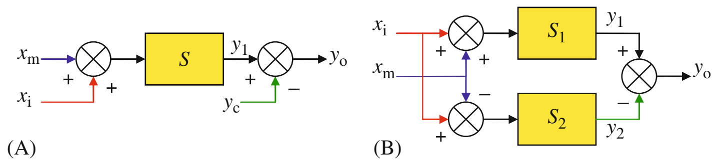

Figure 3.2 General layout of compensation: (A) compensation signal; (B) balanced configuration.

indication of several error sources. In this figure, x_m is the measurement signal and y_o the output signal. *Additive* error signals are modeled as additional input signals: x_d and x_s represent deterministic and stochastic error signals, respectively. They model all kind of interference from the environment and the equivalent error signals due to system offset and noise. The error inputs ε_d and ε_s represent *multiplicative* errors: these signals affect the sensitivity of the sensor. For this simplified model, the output signal of a sensor can be written according to Eq. (3.1):

$$y_o = S(1 + \varepsilon_d + \varepsilon_s)(x_m + x_d + x_s) \tag{3.1}$$

where S is the nominal sensitivity. This model will be used to evaluate various error reduction methods. Some of these methods will reduce mainly additive sensors; others minimize multiplicative errors. Improvement of sensor performance can be obtained through use of a sophisticated design or simply through some additional signal processing. We will discuss five basic error reduction methods:

1. Compensation
2. Feedback
3. Filtering
4. Modulation
5. Correction

The methods not only apply to sensors but also to other signal handling systems as amplifiers and signal transmission systems.

3.2.1 Compensation

Compensation is a simple and effective method to minimize additive errors due to interference signals. The basic idea is as illustrated in Figure 3.2. In Figure 3.2A

The absence of DC and low-frequency components considerably facilitates the amplification of modulated signals: offset, drift and low-frequency noise can be kept far from the new signal frequency band. When very low voltages must be measured, it is recommended to modulate these prior to any other analogue signal processing that might introduce DC errors.

Measurement Bridge as Modulator

The principle of the bridge modulator is illustrated with the resistance measurement bridge or Wheatstone bridge of Figure 3.11.

The bridge is connected to an AC signal source V_i. This AC signal (usually a sine or square wave) acts as the carrier. In this example we consider a bridge with only one resistance (R_3) that is sensitive to the measurand. Assuming equal values of the three other resistances, the signal V_a is just half the carrier, whereas V_b is an AM signal: half the carrier modulated by R_3. The bridge output V_o is the difference between these two signals, so an AM signal with suppressed carrier.

This output can be amplified by a differential amplifier with high gain; its low frequency properties are irrelevant: the only requirements are a sufficiently high bandwidth and a high CMRR for the carrier frequency to accurately amplify the difference $V_a - V_b$.

Modulation techniques also apply to many non-electric signals. An optical signal can be modulated using a LED or laser diode. If the source itself cannot be modulated, optical modulation can be performed by, for instance, a chopping wheel, as is applied in many pyroelectric measurement systems. Also, some magnetic sensors employ the modulation principle. Special cases are discussed in subsequent chapters.

3.2.5 Demodulation

The reverse process of modulation is demodulation (sometimes called detection). Looking at the AM signal with carrier (for instance in Figure 3.7), we observe the similarity between the envelope of the amplitude and the original signal shape. An obvious demodulation method would therefore be envelope detection or peak detection. Clearly, envelope detectors operate only for AM signals *with* carrier. In an AM signal *without* carrier, the envelope is not a copy of the input anymore. Apparently, additional information is required with respect to the phase of the input, for a full recovery of the original waveform.

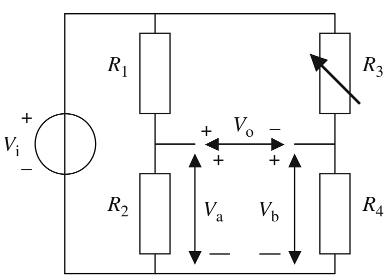

Figure 3.11 Wheatstone bridge as modulator.

Table 4.1 Overview of Resistive Sensor Types

Domain	Measurand	Geometry	Resistivity
Mechanical	Linear displacement	Potentiometer	
	Angle	Potentiometer	
	Strain		Metal strain gauge
	Force, torque, pressure		Piezoresistor
	Acceleration	Potentiometric	
Magnetic	Magnetic field		Magnetoresistor
Thermal	Temperature		Thermistor, Pt100
Optical	Light flux		LDR

Figure 4.1 Various connections to the slider: (A) rotating shaft, (B) sliding rod and (C) flex cable.

Table 4.1 presents an overview of the various resistive sensors discussed in this chapter.

4.2 Potentiometric Sensors

4.2.1 Construction and General Properties

Potentiometric displacement sensors can be divided into linear and angular types, according to their purpose and associated construction. A potentiometric sensor consists of a (linear or toroidal) body which is either wire wound or covered with a conductive film. A slider (or wiper) can move along this conductive body, acting as a movable electrical contact. The connection between the slider and the object of which the displacement should be measured is performed by a rotating shaft (angular potentiometers), a moving rod, an externally accessible slider (sledge type) or a flexible cable that is kept stretched during operation. Figure 4.1 shows a schematic view of some of these constructions. In all cases the resistance wire or film and the wiper contacts should be properly sealed from the environment to minimize mechanical damage and corrosion. This is an important issue when applied in mechatronic systems that operate in harsh environments. Robust potentiometers have a stainless steel shaft or rod, and a housing of, for instance anodized aluminium. The moving parts of the potentiometer are provided with bearings, to

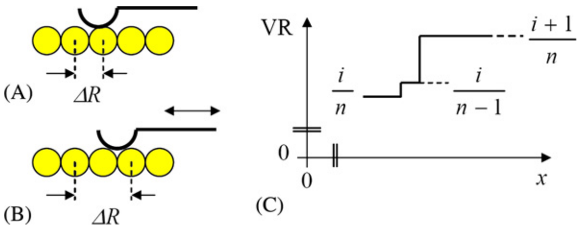

Figure 4.4 (A) Slider on top a winding, (B) slider between two windings and (C) illustration of the resolution of a wire-wound potentiometer.

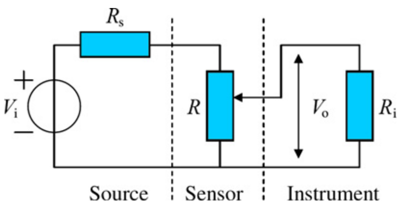

Figure 4.5 Interface circuit for a potentiometer.

can be as good as 0.01% (see Table 4.2). Improper interfacing may introduce additional non-linearity, as will be explained in Section 4.2.3.

The position resolution of a wire-wound potentiometer is set by the number of turns n. With R the total resistance, the resistance of a single turn amounts $\Delta R = R/n$. As the wiper steps from one turn to the next, the VR changes leap-wise with an amount of $1/n$ when the wiper moves continuously (Figure 4.4A); hence the resolution equals $\Delta R/R = 1/n$. At wiper position on top of turn i, VR equals i/n; on top of the next turn it increases to $(i+1)/n$. Actually the wiper may short circuit one turn when positioned just between two windings (Figure 4.4B). In those particular positions the total resistance drops down to $(n-1)\Delta R$, hence $\mathrm{VR} = i/(n-1)$, which is slightly more than i/n as shown in Figure 4.4C.

The resolution can be increased (without change of outer dimension) by reducing the wire thickness. However this degrades the reliability because a thinner wire is less wear resistant. The resolution of a film potentiometer is limited by the size of the carbon or silver grains that are impregnated in the plastic layer to turn it into a conductor. The grain size is about 0.01 μm; the resolution is about 0.1 μm at best.

4.2.3 Interfacing

The interfacing of a potentiometric sensor is essentially simple (Figure 4.5).

To measure the position of the wiper, the sensor is connected to a voltage source V_i with source resistance R_s; the output voltage on the wiper, V_o, is measured by an instrument with input resistance R_i. Ideally, the voltage transfer V_o/V_i equals the VR. Due to the presence of a source resistance and load resistance, the transfer might differ from the VR. We will calculate the error introduced by both these effects.

Strain gauges respond primarily on strain, $\Delta l/l$. Using Hooke's law the applied force is found from the value of the compliance or elasticity of the material on which the strain gauge is fixed.

There are two ways strain gauges are applied in practice:

1. mounted directly on the object whose strain and stress behaviour has to be measured; when cemented properly, the strain of the object is transferred ideally to the strain gauge (for instance to measure the bending of a robot arm);
2. mounted on a specially designed spring element (a bar, ring or yoke) to which the force to be measured can be applied (for instance to measure stress in driving cables and guys).

Strain gauges are excellent devices for the measurement of force and torque in a mechatronic construction. The unbound gauges are small and can be mounted on almost any part of the construction that experiences a mechanical force or moment.

Another approach is to include load cells in the construction. A *load cell* consists of a metal spring element on which strain gauges are cemented. The load is applied to this spring element. The position of the (preferably four) strain gauges on the spring element is chosen such that one pair of gauges is loaded with compressive stress and the other pair with tensile stress (differential mode). If the construction does not allow such a configuration, the two pairs of strain gauges are mounted in a way that one pair experiences the strain that has to be measured, while the other pair is (ideally) not affected by the strain. The second pair merely serves for offset stability and temperature compensation. Obviously, the construction of the spring element and the strain gauge arrangement determine the major properties of the device. Figure 4.13 shows several designs of spring elements for the measurement of force, for various ranges.

Figure 4.13A presents a typical construction for large loads. The central part of the transducer is a metal bar with reduced cross-sectional area where gauges are mounted. There are four strain gauges in a full-bridge configuration. Two strain gauges (front and back side) measure the axial force, with equal sensitivity; two other gauges measure the transverse force, also with equal sensitivity. The gauges

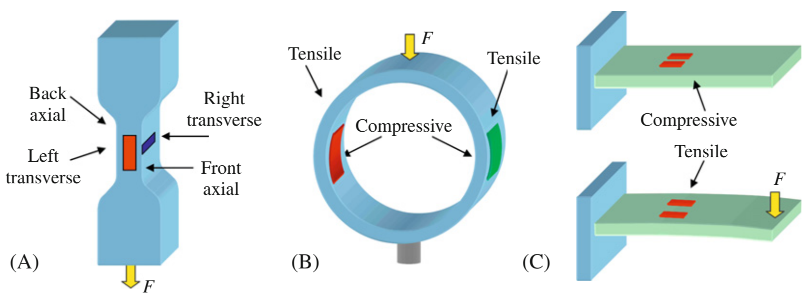

Figure 4.13 Different spring element designs: (A) column type, (B) yoke type and (C) bending beam.

silicon: this material can be used as the carrier of the sensor and, moreover, part of the interface electronics can be integrated with the sensor on the same carrier. Elastomers can be made piezoresistive by a special treatment, for instance by adding conductive particles to the non-conducting elastic material.

Piezoresistive Silicon Sensors

The underlying physical principle of piezoresistivity in silicon goes back to the energy band structure of the silicon atom. An applied mechanical stress will change the band gap. Depending on the direction of the applied force with respect to the crystal orientation, the average mobility of electrons in n-type silicon is reduced, resulting in an increase of the resistivity. So the gauge factor of n-type silicon is negative and reaches values as high as −150. The absolute gauge factor decreases with increasing doping concentration.

In p-silicon, holes are the majority carriers: their mobility is influenced by the position in the valance band. The gauge factor of p-silicon appears to be larger than n-silicon (at the same temperature and doping concentration), and has a positive value. In both cases the piezoresistive effect dominates the geometric effect (as used in metal strain gauges).

Table 4.4 shows numerical values for the gauge factor of p- and n-doped silicon, for three different crystal orientations, and a doping level corresponding to a resistivity of 1 Ω cm. Figure 4.16 displays these three main crystal orientations of silicon.

Table 4.4 Some Gauge Factors for Silicon [12]

Orientation	**p-Type**	**n-Type**
[111]	173	−13
[110]	121	−89
[100]	5	−153

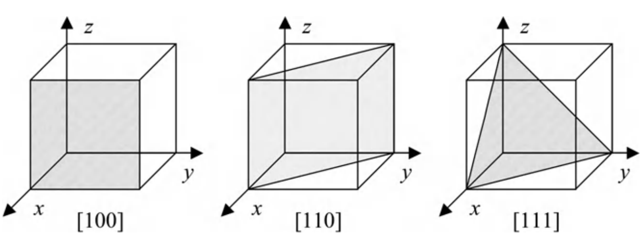

Figure 4.16 Three orientation surfaces in silicon.

Although the resistance change is primarily caused by material deformation, it is common use to express the piezoresistivity of silicon in terms of pressure sensitivity:

$$\frac{\Delta R}{R} = K \cdot S = K \cdot s \cdot T = \pi \cdot T \tag{4.38}$$

with K the gauge factor as defined in Eq. (4.12), π the piezoresistivity (m^2/N) and S and T the mechanical strain and tension in the material, respectively. This is a simplified expression: the piezoresistive coefficient depends strongly on the direction of the applied force relative to the crystal orientation. The pressure sensitivity π of piezoresistive sensors (in silicon) depends on three factors:

1. the conductivity (orientation dependent)
2. the direction of the applied force
3. the orientation of the resistors with respect to the crystal orientation

The second factor in this list is related to the elastic behaviour of the material, as already described by the 6×6 compliance matrix (Appendix A):

$$S_k = s_{kl} T_l, \quad k, l = 1, \ldots, 6 \tag{4.39}$$

The combined orientation-dependent conductivity and compliance yields an expression for the relative resistance change r as a function of the vector T:

$$r_i = \pi_{ij} \cdot T_j, \quad i, j = 1, \ldots, 6 \tag{4.40}$$

For silicon many of the matrix elements are zero due to crystal symmetry, and some are pair-wise equal. This results in the piezoresistivity matrix equation for silicon:

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{pmatrix} \cdot \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{pmatrix} \tag{4.41}$$

So there are only three independent components describing the piezoresistivity of silicon. Their numerical values are given in Table 4.5.

n-Type silicon appears to have a strong negative piezoresistivity in the x-direction, and about half as much positive in the y- and z-directions; p-silicon is less sensitive in these directions. However a shear force (with respect to an arbitrary direction) results in a large resistance change.

4.4.2 Micromachined Piezoresistive Sensors

The unremitting progress in micromachining technology and the creation of microelectromechanical devices (MEMS) have great impact on sensor development. Advantages of this technology are:

- all piezoresistors are deposited in one processing step, resulting in almost identical properties;
- the resistors on the membrane can be configured in a bridge;
- the resistors have nearly the same temperature, due to the high thermal conductivity of silicon;
- interface electronics, including further temperature compensation circuits, can be integrated with the sensor bridge on the same substrate;
- the dimensions of the device can be extremely small; they are mainly set by the package size and the mechanical interface.

Sensors based on this technology are now widely spread, but much research is still going on to further improve the (overall) performance. MEMS also have their limitations, and other technologies are also being investigated and further developed to create better or cheaper sensors. There is also a tendency to combine technologies to benefit from the advantages provided by each of them.

Pressure and Force Sensors

Figure 4.18 shows the basic configurations of a silicon *pressure sensor* and a silicon *force sensor*, both based on piezoresistive silicon.

The sensor carrier or substrate is a silicon chip: a rectangular part cut from the wafer, with thickness about 0.6 mm. The wafer material is lightly doped with positive charge carriers, resulting in p-type silicon. On top of the wafer, a thin layer of n-type silicon is grown, called the epitaxial layer or *epilayer*. The substrate is locally etched away from the bottom up to the epitaxial layer, using selective etching technology. The result is a thin silicon membrane consisting of only the epitaxial layer, the thickness of which is some μm. This membrane acts as a deformable element.

Using standard silicon processing technology, piezoresistive sensors are deposited at positions on the membrane where the deformation is greatest. The gauge factor of silicon is much larger than that of metals; however the temperature

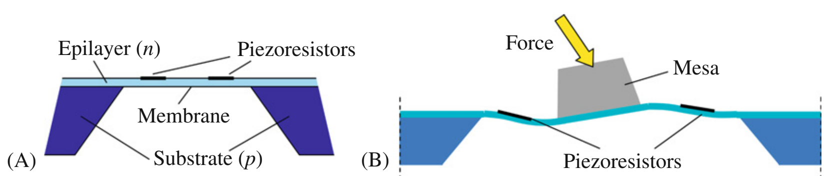

Figure 4.18 Cross section of (A) a piezoresistive pressure sensor in silicon technology (simplified), (B) shear force sensor (not on scale).

Suitable sensor materials are piezoresistive polymers, rubbers and carbon as coating material. Fibres from these materials can be interwoven with the textile. Stretching and bending of the textile result in elongation of the fibres, and hence a change in resistance. An application example of such wearable sensing is given in Ref. [22], reporting about such sensors for the measurement of aspiration. The elongation of the fibre can be as large as 23%, resulting in a resistance change of about 300%.

Piezoresistive sensors are also found in *inclinometers*. Such sensors measure the tilt angle, that is the angle with respect to the earth's normal. Mounted on a robot, for instance, the sensor provides important data about its vertical orientation, which is of particular interest for walking (or legged) robots. In Ref. [23] a micromachined inclinometer was proposed, based on silicon piezoresistive sensors. The sensor consists of a micromass suspended on thin beams. Gravity forces the mass to move towards the earth's centre of gravity. The resulting bending of the beams is measured in two directions, by properly positioned integrated piezoresistors. The authors report an average sensitivity of about 0.1 mV per degree over a range of $\pm 70^\circ$ inclination.

Tactile Sensors

Piezoresistive elastomers can be used for a variety of sensing tasks in robotics. The pressure sensitivity of the bulk resistance is useful to sense touch (recall the high sensitivity for small forces), to measure gripping force and for tactile sensing. The material is shaped in sheets, which is very convenient for the construction of flat sensors, and in particular for tactile sensors. Piezoresistive elastomers belong to the first *tactile sensors* in robot grippers [24,25]. They still receive much attention from designers of robots intended for human-like capabilities, in particular soft gripping (in horticultural applications, for example).

Most resistive tactile sensors are based on some kind of piezoresistive elastomer and are of the cutaneous type. At least one company offers such pressure-sensitive devices, which can also be used for the construction of tactile sensors [26]. Many researchers have reported on the usefulness of these so-called force-sensitive resistors (FSR) as tactile sensors, for instance in Refs [27–29].

Not only is the bulk resistivity a proper sensing parameter, but it is also possible to utilize the surface resistivity of such materials. The contact resistance between two conductive sheets or between one sheet and a conducting layer changes with pressure, mainly due to an increase of the contact area (Figure 4.20). The resistance–pressure characteristic is similar to that of Figure 4.17.

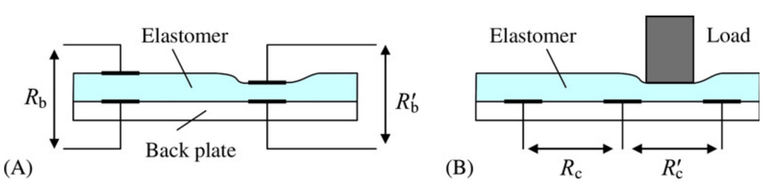

Figure 4.20 Piezoresistive tactile sensors: (A) bulk mode and (B) contact mode.

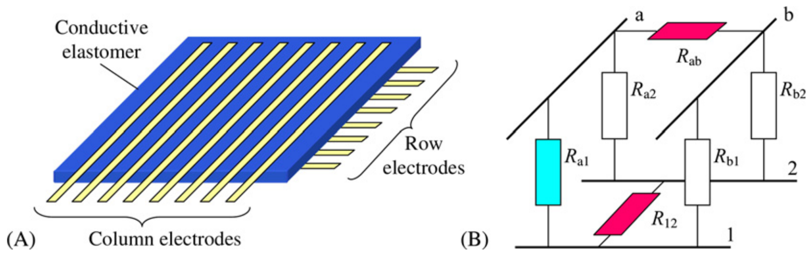

Figure 4.21 Basic idea of row−column readout of a tactile matrix: (A) row and column electrodes on a piezoresistive elastomer and (B) model with four taxels, showing shunt resistances.

An important aspect of tactile sensors based on resistive sheets is the selection and readout of the individual pressure points (taxels). Most popular is the row−column readout, accomplished by a line grid of highly conductive electrodes on either side of the elastomer, making right angles and thus defining the pressure points of the sensor matrix (Figure 4.21A).

Individual taxels are addressed by selecting the corresponding row and column, for instance by applying a reference voltage on the column electrode and measuring the current through the selected row electrode. By using multiplexers for both the row and the column selection, the whole matrix can be scanned quickly. However due to the continuous nature of the resistivity of the sheet, the selected taxel resistance is shunted by resistances of all other taxels, as can be seen in the model of the tactile sensor given in Figure 4.21B. For instance when selecting taxel a-1, the taxel resistance R_{a1} is shunted by resistances $R_{ab} + R_{b1}$ and $R_{a2} + R_{12}$, resulting in unwanted crosstalk between the taxels. Even when the inter-electrode resistances R_{ab} and R_{12} are large compared to the taxel resistances, the selected taxel resistance is shunted by $R_{a2} + R_{2b} + R_{b1}$. As a consequence when one or more of the taxels a-2, b-2 and b-1 are loaded, the unloaded taxel a-1 is virtually loaded. This phenomenon is denoted by 'phantom images'. In an $n \times m$ matrix many of such phantom images are seen by the selected taxel, an effect that is more pronounced when loading multiple taxels of the tactile device.

Crosstalk and phantom images are reduced by actively guarding the non-selected rows and columns, thus zeroing the potential over all non-selected taxels. The principle is illustrated with the simple 2×2 matrix of Figure 4.22.

In Figure 4.22A selection of taxel a-1 is performed by connecting a voltage V_a to column electrode **a**, and measuring the resulting current through row electrode **1**, while all other rows and columns remain floating. Obviously since an additional current component I_2 flows through the other three taxel resistances, the apparent taxel resistance amounts $V_a/I_1 = R_{a1}//(R_{a2} + R_{2b} + R_{b1})$. In Figure 4.22B the non-selected row electrodes are all connected to ground; the additional current, which is now I_2', flows directly to ground, so the measured resistance is $V_a/I_1 = R_{a1}$, which is just the resistance of the selected taxel. Note that the current through the selected column electrode **a** can be quite large, in particular when many taxels on this

[3] M.J. Caruso, L.S. Withanawasam: Vehicle detection and compass applications using AMR magnetic sensors; Company information Honeywell, Plymouth, MN, 1999.
[4] J.E. Lenz: A review of magnetic sensors, Proc. IEEE, 78(6) (1990), 973−989.

Books and Review Articles on Tactile Sensors

[5] H.R. Nicholls (ed.): Advanced tactile sensing for robotics, World Scientific, 1992; ISBN 981-02-0870-7.
[6] R.A. Russell: Robot tactile sensing, 1990; ISBN 0−13-781592-1.
[7] A. Pugh (ed.): Robot sensors, Vol. 2: Tactile and non-vision, 1986; ISBN 0−948507-02-0.
[8] L.D Harmon: Automated tactile sensing, Int. J. Rob. Res., 1(2) (1982), 3−32.

5 Capacitive Sensors

Capacitive sensors for displacement and force measurements have a number of advantages. A capacitor consists of a pair of conductors; since no other materials are involved, capacitive sensors are very robust and stable and applicable at high temperatures and in harsh environments. The dimensions of capacitive sensors may vary from extremely small (in MEMS) up to very large (several metres). The theoretical relation between displacement and capacitance is governed by a simple expression, which in practice can be approximated with high accuracy, resulting in a very high linearity. Using special constructions, the measurement range of capacitive sensors can be extended almost without limit while maintaining the intrinsic accuracy. Moreover, because of the analogue nature of the capacitive principle, the sensors have excellent resolution.

This chapter starts with resuming the notions capacitance and permittivity. Next, we discuss basic configurations for capacitive sensors. The capacitance value changes with variation in the geometry. We will focus in particular on linear and angular displacement sensors and on force sensors. Next, integrated silicon capacitive sensors are briefly reviewed and some interface circuits are given to measure small capacitance changes.

5.1 Capacitance and Permittivity

The capacitance (or capacity) of an isolated conducting body is defined as $Q = C \cdot V$, where Q is the charge on the conductor and V the potential (relative to 'infinity', where the potential is zero by definition). To put it differently, when a charge Q from infinite distance is transferred to the conductor, its potential becomes $V = Q/C$.

In practice we have a set of conductors instead of just one conductor. When a charge Q is transferred from one conductor to another, the result is a voltage difference V equal to Q/C; the conductors are oppositely charged with Q and $-Q$, respectively (Figure 5.1). Again, for this pair of conductors, $Q = C \cdot V$, where V is the voltage difference between these conductors.

Usually, the capacitance is defined 'between' two conductors: this capacitance is determined exclusively by the geometry of the complete set of conductors and the dielectric properties of the (non-conducting) matter in between the conductors and not by the potential of the conductors. The set of conductors is called a capacitor.

Sensors for Mechatronics. DOI: 10.1016/B978-0-12-391497-2.00005-4

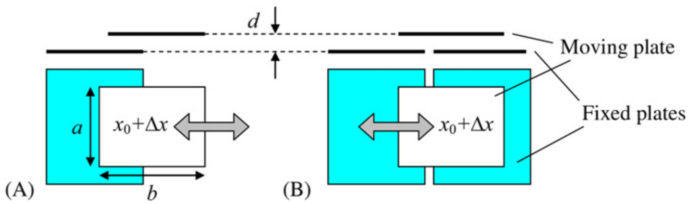

Figure 5.5 Basic capacitive displacement sensors with variable surface area: (A) single mode and (B) differential mode.

deformable membrane, pressure can also be measured using the capacitive principle. Generally, electric fields are better manageable than magnetic fields: by (active) guarding, it is easy to create electric fields that are homogeneous over a wide area. This is the major reason that displacement sensors based on capacitive principles have excellent linearity.

Figure 5.5 shows, schematically, two basic configurations for linear displacement. The moving object of which the displacement has to be measured is connected to the upper plate. In these examples the parameter A (effective surface area) varies with displacement.

In Figure 5.5A a linear displacement of the upper plate in the indicated direction introduces a capacitance change which is, ideally, $\Delta C = \varepsilon \Delta x a/d$, or a relative change $\Delta C/C = \Delta x/x$. All other parameters need to be constant during movement of the plate. Capacitive sensors for rotation can be configured in a similar way, with segmented plates rotating relatively to each other, where the distance d remains constant. To assure a linear relation between the displacement and the capacitance change, active guarding has to be applied in all cases.

Figure 5.5B shows a differential transducer (see Chapter 3): a displacement of the upper plate causes two capacitances to change simultaneously but with opposed sign: $\Delta C_1 = -\Delta C_2 = \varepsilon \Delta x a/d$. The initial position is defined as the position for which $C_1 = C_2$. For this position ($\Delta x = 0$) a change in a or d (due to for instance play, backlash or temperature change) does not introduce a zero error, as long as both capacitance changes are equal.

An additional advantage of the differential configuration is the extended dynamic range: in the reference or initial position the capacitances of the two capacitors are equal. Since only the difference is processed, the initial output signal is zero. A small displacement results in a small output signal that can be electronically amplified without overload problems. A single capacitance in the initial position may produce a considerable output signal, which cannot be amplified much more, unless it is first compensated by an equal but opposite offset signal. However, compensation by an equally shaped sensor is more stable than compensation in the electrical domain.

A disadvantage of the configurations shown in Figure 5.5 is the electrical connection to the moving plate, required for the supply and transmission of the measurement signals. Figure 5.6 shows a configuration without the need for such a connection.

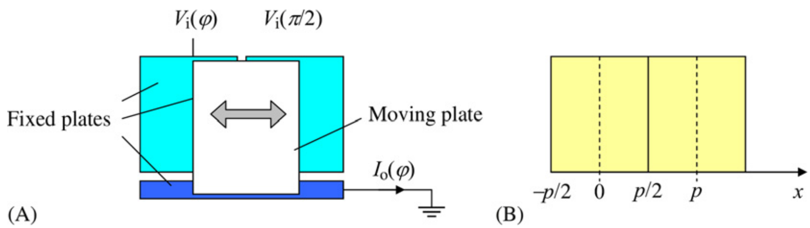

Figure 5.7 Phase read-out of a capacitive displacement sensor: (A) structure and (B) co-ordinates.

To calculate the phase of the output current in dependence of the displacement we define the co-ordinates as in Figure 5.7B. The centre of the moving plate runs from $x = 0$ to $x = p$. For simplicity all plates have equal width p. Suppose the input voltages are $V_1 = \hat{V}\sin\omega t$ and $V_2 = \hat{V}\cos\omega t$. According to Eq. (5.10) the output current equals

$$I_o(t) = \frac{C_1\omega\hat{V}\cos\omega t + C_2\omega\hat{V}\sin\omega t}{1 + C_1/C_3 + C_2/C_3} \tag{5.12}$$

The capacitances C_1 and C_2 have the value $\varepsilon aw/d = c'w$, with a the (constant) plate length and w the effective width, running from p to 0 for C_1 and from 0 to p for C_2. With Figure 5.7B the capacitances are found:

$$\begin{aligned} C_1(x) &= c'(p - x) \\ C_2(x) &= c'x \end{aligned} \tag{5.13}$$

Substitution in Eq. (5.12) results in

$$I_o(t) = K\frac{(p - x)\cos\omega t + x\sin\omega t}{1 + c'p/C_3} = \hat{I}\cos(\omega t + \varphi) \tag{5.14}$$

where K is a constant. The phase angle satisfies the relation

$$\varphi = \arctan\frac{x}{p - x} \tag{5.15}$$

The phase varies with displacement in a slightly non-linear way (Figure 5.8). Stray fields cause further deviations from linearity, but even with guarding the relation is essentially nonlinear. However, by modifying the electrode shapes it is possible to get an almost linear relationship.

Capacitive displacement sensors can also be configured in a cylindrical configuration, as in Figure 5.3C. The principle is the same, but a cylindrical set-up is more compact, has less stray capacitances and therefore a better linearity. This type of capacitive sensor is called a linear variable differential capacitor (LVDC). The sensor exhibits extremely good linearity (better than 0.01%) and a low temperature sensitivity (down to 10 ppm/K).

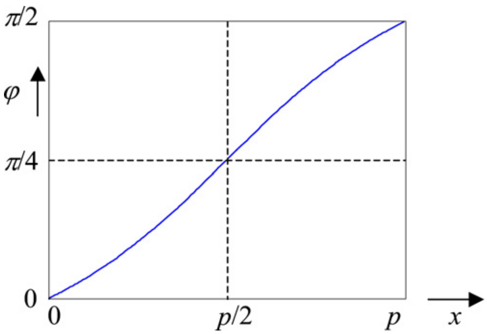

Figure 5.8 Phase of the output current versus displacement.

The rotational version of the LVDC is called RVDC (rotational variable differential capacitor). It is based on the same principle but configured for angular displacements. Angular sensitivity is obtained using triangularly shaped electrodes. Further specifications on these sensors are given in the overview table (Table 5.2) in Section 5.4.

The lower limits of linearity, stability and accuracy of these precision sensors are set by the air humidity: water vapour in the air may condensate in edges and small gaps of the construction (capillary condensation), thereby locally replacing air (with dielectric constant of about 1) by liquid water (with $\varepsilon_r \approx 80$, see Table 5.1). Even at temperatures above dew point, water vapour may condensate due to contamination: hygroscopic particles attract water molecules and act as condensation nuclei. This means that for the highest performance the temperature of the sensor should be kept well above the dew-point temperature.

5.2.2 Multiplate Capacitive Sensors

The measurement range of the linear capacitive sensors discussed so far is limited to approximately the width of the moving plate (Figures 5.5–5.7). However, the range can simply be extended by repeating the basic structure of the previously discussed configuration. Figure 5.9A shows this concept [8]. It consists of an array of sections similar to the one in Figure 5.7. When the moving plate approaches the end of one section, it simultaneously enters the next section.

The fixed electrodes in the array are connected to sine wave voltages with phase differences of $\pi/2$. When the moving plate travels along four successive sections, the phase of the voltage on this plate changes from 0 to 2π (Figure 5.9B). This is repeated for each group of four electrodes. Hence the output phase changes periodically with displacement, with a period of four times the plate pitch p. An unambiguous output is obtained by keeping track of the number of passed cycles, using an incremental counter.

The straight lines in Figure 5.9B are actually four s-shaped curves from Figure 5.8 in series, but nonlinearity can be compensated for by appropriate shaping of the fixed electrodes. Phase differences can be measured with high resolution, down to 0.1°. By periodically repeating the structure a large range is obtained

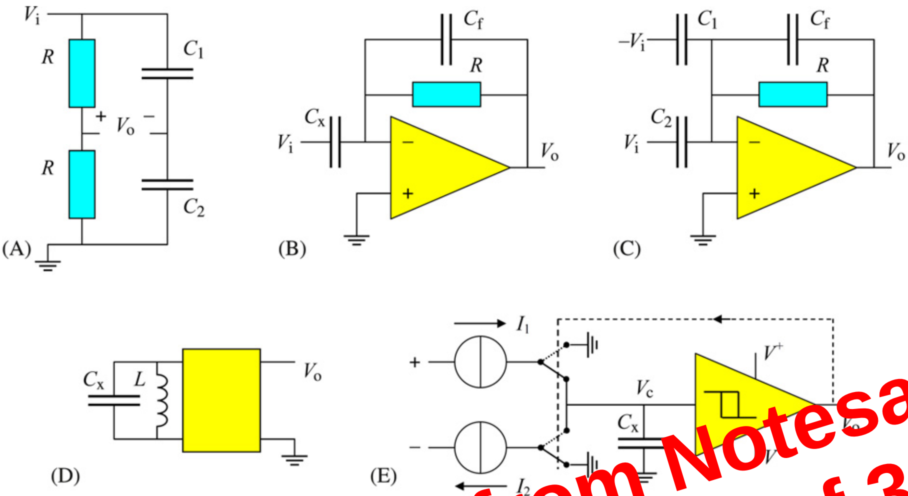

Figure 5.13 Methods for the measurement of capacitance changes: (A) bridge; (B) and (C) current–voltage measurement—single and differential; (D) frequency measurement; (E) time measurement.

The transfer is frequency independent but nonlinear. In a differential or balanced configuration $C_1 = C + \Delta C$ and $C_2 = C - \Delta C$, thus

$$\frac{V_o}{V_i} = -\frac{\Delta C}{2C} \tag{5.18}$$

Clearly, the differential mode yields better linearity and a wider dynamic range (see also Chapter 4 on resistance bridges). A proper measurement of the bridge output voltage requires a differential amplifier with high input impedance and high common mode rejection ratio.

Interface circuits based on current–voltage measurements are shown in Figure 5.13B and C: the first for single mode operation and the second for differential mode operation. If resistance R is disregarded, the transfer function of the single mode circuit is

$$\frac{V_o}{V_i} = \frac{C + \Delta C}{C_f} \tag{5.19}$$

and for the differential mode:

$$\frac{V_o}{V_i} = 2\frac{\Delta C}{C_f} \tag{5.20}$$

A totally different application of a capacitive sensor is found in Ref. [39]: the measurement of tyre *strain*. Car tyres are usually reinforced with layers of steel wires. When the tyre is loaded, the resulting strain causes the distance between these wires to increase. Hence, the capacitance between the wires changes accordingly. The paper describes how to measure this capacitance change. The interface is based on an oscillator (like in Figure 5.13D), but instead of an *LC*-oscillator, an *RC*-oscillator is chosen. Wireless read-out allows strain measurement during rotation of the wheel with tyre. Another approach to measure tyre strain using a capacitive method is described in Ref. [40]. The authors have fabricated a particular strain gauge, consisting of a pair of interdigitated flat electrodes on a flexible polyimide carrier. The gauge, glued onto the inside tyre, responds to strain by a change in capacitance [41].

As mentioned before, capacitive sensors can also be applied for *force* and *pressure* measurements. A particular application is on-line weighing of (for instance) cars [42]. The sensor consists of two rubber layers sandwiched between (three) conductive sheets. When pressed, the distance between the conductive sheets becomes smaller, producing an increase in capacitance. The sensor is flexible, light-weight and easy to carry. It can simply be positioned on the road and is capable of dynamically measuring the weight of a (slowly) passing car.

Thus far we have used the geometric factor G in Eq. (5.5) as the basis for a capacitive measurement. The dielectric constant ε_r is another parameter determining the capacitance value of a capacitive structure and hence a workable method for sensing. Measuring liquid *level* in a tank is an obvious example. A pair of flat-plate electrodes hanging vertically in a tank represents a capacitance that depends on the dielectric constant of the fluid between these plates (Figure 5.15A).

In an empty tank this is air (or another gas with $\varepsilon_r \approx 1$), in a full tank the liquid (for instance water with $\varepsilon_r \approx 80$). A partly filled capacitor can be conceived as two capacitors in parallel: one filled with the liquid up to the upper level (C_L), and one with air from the level to the top of the plates (C_A). The total capacitance is the sum of these two capacitances, and hence a measure for the level. Obviously, this method holds only for non-conducting liquids. Since water is both dielectric and conductive, at least one of the electrodes should be isolated from the liquid, and

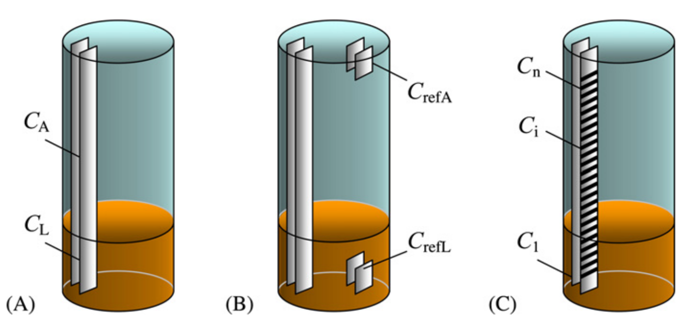

Figure 5.15 Capacitive level measurements.

Table 6.2 Comparison Between the Electrical and the Magnetic Domain

Electrical Domain	Magnetic Domain	
$E = \frac{1}{\sigma} \cdot J$	$H = \frac{1}{\mu} \cdot B$	(6.7)
$V = \int E \cdot \mathrm{d}l$	$n \cdot I = \int H \cdot \mathrm{d}l$	(6.8)
$I = \iint J \cdot \mathrm{d}A$	$\Phi = \iint B \cdot \mathrm{d}A$	(6.9)
$V = R_\mathrm{e} \cdot I$	$n \cdot I = R_\mathrm{m} \cdot \Phi$	(6.10)
$R_\mathrm{e} = \frac{1}{\sigma} \cdot \frac{l}{A}$	$R_\mathrm{m} = \frac{1}{\mu} \cdot \frac{l}{A}$	(6.11)

elements is the same. So the resistances (reluctances) of these elements can simply be summed to find the total reluctance of the series circuit.

The *self-inductance* of a magnetic circuit with coupled flux is found as follows. The induced voltage equals $V_\mathrm{ind} = n \cdot (\mathrm{d}\Phi/\mathrm{d}t)$ (when there are n turns). Substitution of Φ using Eq. (6.9) yields $V_\mathrm{ind} = (n^2/R_\mathrm{m}) \cdot (\mathrm{d}I/\mathrm{d}t)$ and since $V = L(\mathrm{d}I/\mathrm{d}t)$ the self-inductance is:

$$L = \frac{n^2}{R_\mathrm{m}} = n^2 \cdot \frac{\mu A}{l} \tag{6.12}$$

So the coefficient of self-inductance (unit Henry, H = Wb/A) is proportional to the square of the number of turns and inversely proportional to the reluctance. Several sensors, based on a change in self-inductance and reluctance, will be further discussed in this chapter.

6.1.5 Magnetostriction

All ferromagnetic materials exhibit the magnetostrictive effect. Basically it is the change in outer dimensions of the material when subjected to an external magnetic field. In the absence of an external field the magnetic domains (elementary magnetic dipoles) are randomly oriented. When a magnetic field is applied, these domains tend to line up with the field, up to the point of saturation. The effect is not strong: materials with a large magnetostriction (for instance Terfenol-D) show

6.2.1 Coil

The most straightforward method for the transduction from magnetic field to an electric voltage is a coil: Eq. (6.5) relates the induced voltage in a coil to the magnetic flux. At first sight, only AC fields can be measured in this way since the induced voltage is proportional to the rate of change in flux. Static fields can nevertheless be measured, just by rotating the coil. Let the surface area of the coil be A and the frequency of rotation ω, then for a homogeneous induction field B, the induced voltage equals:

$$V(t) = -B \cdot A \cdot \frac{\mathrm{d} \sin \alpha(t)}{\mathrm{d}t} = -B \cdot A \cdot \omega \cdot \cos \omega t \tag{6.15}$$

With a rotating coil very small induction fields can be measured. Disadvantages of the method are movable parts, the need for brushes to make electrical connection to the rotating coil and for an actuator to procure rotation.

6.2.2 Hall Plate Sensors

The Hall plate is based on the magnetoresistive effect. In 1856 W. Thomson (Lord Kelvin) discovered that a magnetic field influences the resistivity of a current-conducting wire (see also Section 4.5 on magnetoresistive sensors). Later this effect was named the Gauss effect. Only after the discovery of the Hall effect in 1879, by the American physicist E.F. Hall, could the Gauss effect be explained. Both the Gauss and the Hall effects are remarkably stronger in semiconductors, so they became important for measurement science only after the development of semiconductor technology.

The Hall effect is caused by the Lorentz forces on moving charge carriers in a solid conductor or semiconductor, when placed in a magnetic field (Figure 6.3). The force F_1 on a particle with charge q and velocity v equals:

$$F_1 = q(v \times B) \tag{6.16}$$

The direction of this force is perpendicular to both B and v (right-hand rule). As a result the flow of charges in the material is deflected and an electric field E is built up, perpendicular to both I and B. The charge carriers experience an electric

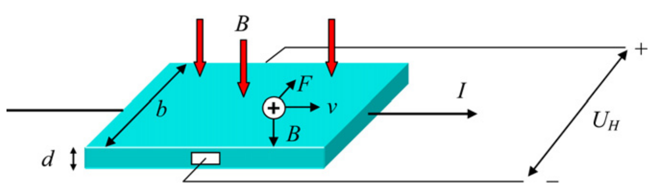

Figure 6.3 Principle of the Hall sensor.

6.3.2 Inductive Proximity and Displacement Sensors

The general principle of a displacement sensor based on variable self-inductance is depicted in Figure 6.7A. The self-inductance of the configuration is approximately equal to $L = n^2/R_m$, according to Eq. (6.12). A displacement Δx of the movable part causes the reluctance R_m to change because the air gap width changes by an amount of $2\Delta x$. With Eq. (6.12), and noting that the iron and the air gaps are in series and hence their reluctances add, the self-inductance modifies according to

$$L(x) = \frac{\mu_0 A n^2}{l_r/\mu_r + 2(x_0 + \Delta x)} = \frac{\mu_0 A n^2}{l_0 + 2\Delta x} = \frac{L_0}{1 + 2\Delta x/l_0} \tag{6.20}$$

where μ_r is the permeability of the iron core, l_r the path length of the flux through the iron part, x_0 the initial air gap, L_0 the self-inductance at the reference position x_0 and $l_0 = l_r/\mu_r + 2x_0$ the total effective flux path in the initial position. Then at a displacement Δx the self-inductance changes by

$$\Delta L = -L_0 \frac{2\Delta x/l_0}{1 + 2\Delta x/l_0} = -L_0 \frac{2\Delta x}{l_0}\left\{1 - \frac{2\Delta x}{l_0} + \frac{1}{2}\left(\frac{2\Delta x}{l_0}\right)^2 - \cdots\right\} \tag{6.21}$$

This result expresses a strongly non-linear sensitivity. A differential configuration gives some improvement (Figure 6.7B). The sensitivity is doubled, whereas the non-linearity is reduced to a second-order effect, as is shown in Eq. (6.22):

$$\begin{aligned} \Delta L = L_2 - L_1 &= \frac{L_0}{1 - 2\Delta x/l_0} - \frac{L_0}{1 + 2\Delta x/l_0} = L_0 \frac{4\Delta x/l_0}{1 - (2\Delta x/l_0)^2} \\ &= L_0 \frac{4\Delta x}{l_0}\left\{1 + \left(\frac{2\Delta x}{l_0}\right)^2 - \cdots\right\} \end{aligned} \tag{6.22}$$

The range of this sensor is limited to $2x_0$. Another variable inductance configuration with a much wider range is shown in Figure 6.8A: a coil with movable core. Since the self-induction varies non-linearly with displacement too, the differential configuration in Figure 6.8B is better because quadratic terms cancel, as was demonstrated in the previous equations.

A common disadvantage of all self-inductance types is the coil itself. The device is not an ideal self-inductance: the wire resistance and the capacitance between the

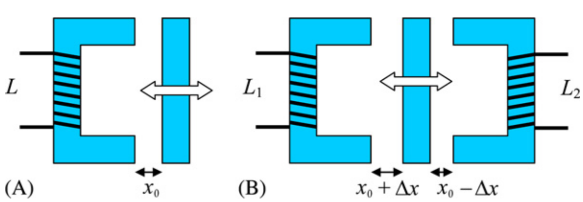

Figure 6.7 Displacement sensors based on variable self-inductance: (A) single and (B) differential.

Another method to detect defects in conductive structures is magnetic flux leakage (MFL). The principle of this technique is shown in Figure 6.22. The object under test is magnetized by an external magnet. In the ideal case the magnetic field is mainly located inside the object because of its high permeability compared to the surrounding air. In case of a surface defect (e.g. hole and crack) the permeability at that location is low, and this causes the magnetic flux to bulge out of the material. This 'leakage' of the magnetic field can be measured by a magnetic sensor located just above the object's surface. Note that the defects create field lines perpendicular to the surface, so sensing in this direction is recommended. With sensitive magnetic sensors it is possible to detect defects in the order of micrometers, and even defects at the bottom side of (relatively thin, several mm) plates or pipe walls.

The MFL technique is still widely studied, aiming at better resolution and a more accurate localization and characterization of defects. A promising technique is the pulsed method that employs the dynamic behaviour of the pulse response for the interpretation of the defects. For instance in Ref. [71] it is shown that surface and sub-surface defects give distinctive responses, and from the measured time-to-peak the depth of the defects can be reconstructed.

A completely different application of a magnetic field sensor is found in Ref. [72]. Goal is to create an image of steel objects embedded in concrete structures. An excitation coil supplied by a DC current generates a static magnetic field over the area of interest. The pattern of the magnetic field reflects the shape of an embedded steel object. An image is obtained by a 2D scanning sequence over the test area by a sensitive magnetic sensor. An optimized layout of the sensor setup and advanced signal processing are required to obtain an image of reasonable quality, but the method enables the visualization of objects at a depth up to10 cm.

6.4.4 Applications of Other Inductive Sensors

Velocity and *acceleration* can be derived from position by taking the first and second derivative of the displacement signal. Real accelerometers make use of a seismic mass: when accelerated, the inertial force causes a deformation of an elastic element connected to the mass. The *accelerometer* described in Ref. [73] is based on an LVDT with a magnetic fluid as movable core. The position of the

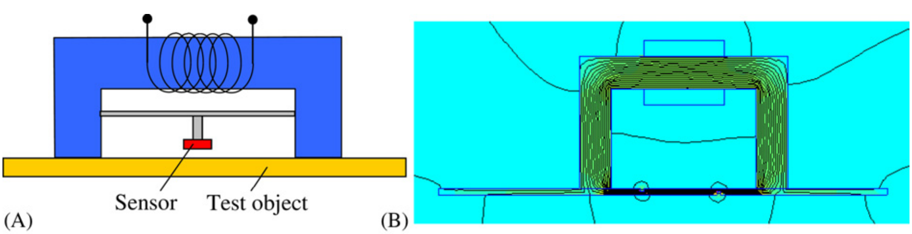

Figure 6.22 Principle of magnetic flux leakage for NDT: (A) measurement setup and (B) simulation example with two defects in the plate under test: left at the top side, right at the bottom side.

[24] A. Baschirotto, E. Dallago, P. Malcovati, M. Marchesi, G. Venchi: A fluxgate magnetic sensor: from PCB to micro-integrated technology, IEEE Trans. Instrum. Meas., 56(1) (2007), 25−31.

[25] B. Andò, S. Baglio, V. Sacco, A.R. Bulsara, V. In: PCB fluxgate magnetometers with a residence times difference readout strategy: the effects of noise, IEEE Trans. Instrum. Meas., 57(1) (2008), 19−24.

[26] H.S. Park, J.S. Hwang, W.Y. Choi, D.S. Shim, K.W. Na, S.O. Choi: Development of micro-fluxgate sensors with electroplated magnetic cores for electronic compass, Sens. Actuators A, 114 (2004), 224−229.

[27] S. Moskowicz: Fluxgate sensor with a special ring-core; Proceedings XVIII IMEKO World Congress, Rio de Janeiro, Brazil, 17−22 September 2006.

[28] H. Grüger, R. Gottfried-Gottfried: Performance and applications of a two axes fluxgate magnetic field sensor fabricated by a CMOS process, Sens. Actuators A, 92 (2001), 61−64.

[29] R.A. Rub, S. Gupta, C.H. Ahn: High directional sensitivity of micromachined magnetic fluxgate sensors; Proceedings Eurosensors XV, Munich, Germany, 10−14 June 2001.

[30] P. Ripka, S.O. Choi, A. Tipek, S. Kawahito, M. Ishida: Symmetrical core improves micro-fluxgate sensors, Sens. Actuators A, 92 (2001), 30−36.

[31] P. Ripka, S. Kawahito, S.O. Choi, A. Tipek, M. Ishida: Micro-fluxgate sensor with closed core, Sens. Actuators A, 91 (2001), 65−69.

[32] P.M. Drlača, P. Kejík, F. Vincent, D. Piguet, F. Gueissaz, R.S. Popović: Single core fully integrated CMOS micro-fluxgate magnetometer, Sens. Actuators A, 110 (2004), 236−241.

[33] S. Kawahito, A. Cerman, K. Aramaki, Y. Tadokoro: A weak magnetic field measurement system using micro-fluxgate sensors and delta−sigma interface, IEEE Trans. Instrum. Meas., 52(1) (2003), 103−110.

[34] S.C. Tang, M.C. Duffy, P. Ripka, W.G. Hurley: Excitation circuit for fluxgate sensor using saturable inductor, Sens. Actuators A, 113 (2004), 156−165.

[35] D. Vyroubal: Impedance of the eddy-current displacement probe: the transformer model, IEEE Trans. Instrum Meas., 53(2) (2004), 384−391.

[36] M.T. Restivo: A case study of induced eddy currents, Sens. Actuators A, 51 (1996), 203−210.

[37] Ph.A. Passeraub, G. Rey-Mermet, P.A. Besse, H. Lorenz, R.S. Popovic: Inductive proximity sensor with a flat coil and a new differential relaxation oscillator; Eurosensors X, Leuven, Belgium, 8−11 September 1996, pp. 375−337.

[38] H. Fenniri, A. Moineau, G. Delaunay: Profile imagery using a flat eddy-current proximity sensor, Sens. Actuators A, 45 (1994), 183−190.

[39] C. Bartoletti, R. Buonanni, L.G. Fantasia, R. Frulla, W. Gaggioli, G. Sacerdoti: The design of a proximity inductive sensor, Meas. Sci. Technol., 9 (1998), 1180−1190.

[40] O. Dezuari, E. Belloy, S.E. Gilbert, M.A.M. Gijs: Printed circuit board integrated fluxgate sensor, Sens. Actuators A, 81 (2000), 200−203.

[41] Y. Hamasaki, T. Ide: A multilayer eddy current microsensor for non-destructive inspection of small diameter pipes; Transducers ’95 − Eurosensors IX, Stockholm, Sweden, 25−29 June 1995, pp. 136−139.

[42] P. Kejík, C. Kluser, R. Bischofberger, R.S. Popović: A low-cost inductive proximity sensor for industrial applications, Sens. Actuators A, 110 (2004), 93−97.

[43] Ph.A. Passeraub, P.A. Besse, S. Hediger, Ch. De Raad, R.S. Popovic: High-resolution miniaturized inductive proximity sensor: characterization and application for step-motor control, Sens. Actuators A, 68 (1998), 257−262.

Sensitivity characteristic: The photo current is proportional to the incident light over more than six decades (Figure 7.3B). This is the main reason why the current mode (or conductivity mode) is preferred over the photovoltaic mode of operation. Current readout is easily accomplished using an operational amplifier in current-to-voltage converter arrangement (Section 7.3.1). Usually a reverse voltage is applied to ensure the diode always operates in the flat part of the characteristics in Figure 7.3A.

Dark current: Not only light (optical energy) but also thermal energy gives rise to the generation of additional charge carriers around the pn-junction. This is exactly why the reverse current of a pn-diode is not zero. Roughly the thermally induced leakage current doubles each 7°C. At room temperature the leakage current of a silicon diode is very small, some 10 pA, but since a photo diode has an essentially larger active area, its dark current is also much larger, and moreover, it increases exponentially with temperature (Figure 7.3C).

Spectral sensitivity: Assuming a quantum efficiency of 100%, each incident photon creates one electron–hole pair, if the photon energy is larger than the bandgap energy E_g (to generate a free electron):

$$E_g = h \cdot \upsilon \tag{7.1}$$

where υ is the frequency and h Planck's constant. So E_g of the material determines the maximum wavelength. Silicon has a bandgap equal to 1.12 eV, corresponding to a limiting wavelength of about 1100 nm. The lower boundary of the spectrum is set by the absorption of the light before it has reached the pn-junction. In general the absorption coefficient of the materials involved grow with reduced wavelength. The result is a spectral band that strongly depends on the materials. Figure 7.3D gives the spectral sensitivities of three photo diodes: Si, GaAs (both sensitive in the visible part of the spectrum) and a special diode for large wavelengths. Detailed information on device characteristics and performance of other types can be found in data books of the manufacturers.

7.1.3 Position Sensitive Diode

A PSD is a light sensitive diode which is not only sensitive to the intensity of the incident light but also to the position where the incoming light beam hits the diode surface. Figure 7.4 shows, schematically, the configuration of a lateral silicon PSD.

The device consists of an elongated pn-photo diode with two connections at its extremities and one common connection at the substrate. An incident light beam can penetrate through the top layer, down to the depleted region around the pn-junction where it generates electron–hole pairs. As in a normal photo diode, the junction is reversed biased. Due to the electric field across the junction, the electrons are driven to the positive side, the holes move to the negative side of the junction. The electrons arriving in the n-region flow through the common contact to ground; the holes in the p-layer split up between the two upper contacts, producing two external photo currents I_a and I_b.

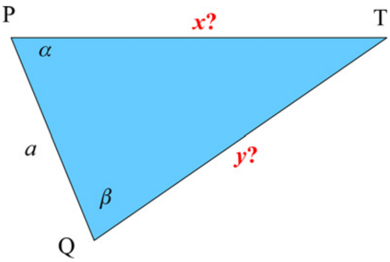

Figure 7.8 Triangulation principle with a PSD.

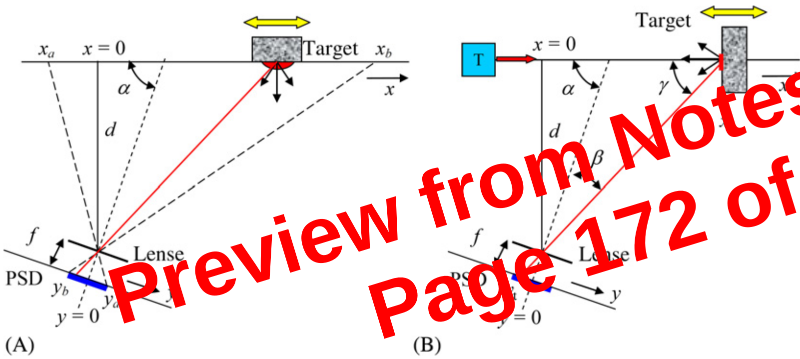

Figure 7.9 Triangulation applied to contact-free distance measurement: (A) direct mode and (B) indirect mode.

x_a to x_b) depends on the size of the PSD (from y_a to y_b), and the focal length f and can further be adjusted with the parameters d and α.

Figure 7.9B gives a configuration in indirect mode. A light beam from a laser source casts a spot on a suitable, well-defined point at the movable target. Part of the scattered light is projected onto the PSD. A displacement of the target causes a displacement of the projected light spot on the sensor. The x-position of the laser source is irrelevant here, as long as the spot intensity is large enough to create a bright spot on the PSD. Clearly each target position x_t corresponds to a spot position y_t on the PSD. This relation depends on the geometry and the focal distance of the lens and can be found using the goniometric relation:

$$\tan\alpha = \tan(\beta + \gamma) = \frac{\tan\beta + \tan\gamma}{1 - \tan\beta\tan\gamma} \tag{7.9}$$

With $y/f = \tan\beta$ and $a/x = \tan\gamma$ (see Figure 7.9B) we arrive at

$$y = f \cdot \frac{x\tan\alpha - d}{x + d\tan\alpha} \quad \text{or} \quad x = d \cdot \frac{y\tan\alpha + f}{f\tan\alpha - y} \tag{7.10}$$

Table 7.3 Comparison Between PSD and Camera for Position Measurements

Item	**PSD**	**Camera (b/w)**
Position information	Analogue	Quantized
Number of positions	1	#pixels
Position resolution	1 ppm	4096 pixels
Dynamic position range	$1{:}10^6$	$1{:}10^3$
Accuracy	High	Calibration required
Speed	30 kHz per point	80 MHz pixel rate (20 kHz line rate)
Light source	LED, laser	Environmental

presents some basic differences between these two devices which could be helpful in making a proper choice. Note that numerical data are typical; higher values for resolution and speed are available.

7.2.3 Optical Encoders

The optical sensors described in this section are digital in nature. They convert, through an optical intermediate, the measurement quantity into a binary signal, representing a binary coded measurement value. Sensor types that belong to this category are optical encoders, designed for measuring linear and angular displacement, optical tachometers (measuring angular speed or the number of revolutions per unit of time) and optical bar code systems for identification purposes.

Optical encoders are composed of a light source, a light sensor and a coding device (much the same as the general setup in Figure 7.1). The coding device consists of a flat strip for linear displacement or disc for angular displacement, containing a pattern of alternating opaque and transparent segments (the transmission mode) or alternating reflective and absorbing segments (the reflection mode). Both cases are illustrated in Figure 7.11.

The coding device can move relatively to the assembly of transmitter and receiver causing the radiant transfer between them switch between a high and a low value. In the transmission mode the encoder consists either of a translucent material (e.g. glass, plastic and mylar), covered with a pattern of an opaque material (for instance a metallization), or just the reverse, for instance a metal plate with slots or holes.

Two basic encoder types are distinguished: absolute and incremental encoders. An *absolute encoder* gives instantaneous information about the absolute displacement or the angular position. Figure 7.12 gives examples of absolute encoder devices in transmission mode.

Each (discrete) position corresponds to a unique code, which is obtained by an optical readout system that is basically a multiple version of Figure 7.11A. The acquisition of absolute position with a resolution of n bits requires at least n optical tracks on the encoder and n separate sensors.

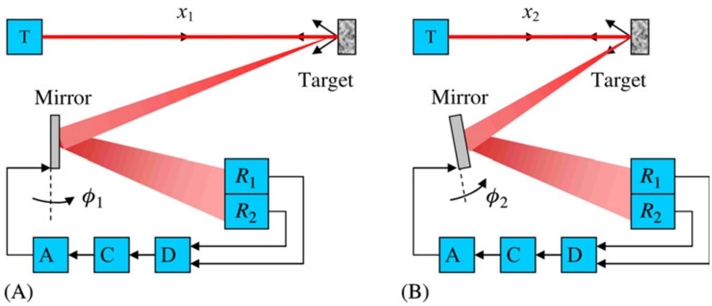

Figure 7.28 Optical position measurement in feedback configuration; (A) far and (B) close by.

To eliminate the influence of changes in source intensity and other unwanted interference in an optical displacement system based on intensity modulation, multiple sensing and feedback are effective solutions. Adding one or more receivers to compensate for the effect of unwanted changes makes reflection-based systems less sensitive to the surface properties and may even allow uncalibrated operation [25,26]. Another solution is presented in Ref. [27], where the relative position of transmitter and receiver is adjusted such that the maximum intensity is received. The distance is found irrespective of the surface characteristics. Feedback is an even more powerful solution but has higher system complexity. Figure 7.28 shows an example of such a system.

The light reflected by the (movable) target is projected onto two photo-sensors via a rotating mirror. The position of this mirror is controlled such that the two sensors receive equal radiation (position x_1 in Figure 7.28A). The corresponding angle of the mirror is ϕ_1. Upon displacement of the object to position x_2 the light beam shifts from the centre of the dual sensor, resulting in different photo currents. This difference is used to drive the mirror to rotate to an angle ϕ_2 where the photo currents become equal again. The control signal or the angle is a measure for the displacement. By feedback the requirements to the source and the sensors are reduced significantly, as explained in Chapter 2. For 2D displacements the moving mirror has 2 d.o.f. and the detector comprises four receivers arranged in a square (a quad).

The principle is applied in for instance *automatic focussing* systems in certain camera types. Another application is found in contact-free measurement of *profiles*. Commercial instruments for that purpose are available with a resolution down to 50 nm, and a range of about 1 m.

All aforementioned examples use the change in irradiance with distance between source and sensor (usually a $1/x^2$ relation). In a light absorbing (or attenuating) medium the intensity drops exponentially, according to Lambert–Beer's law:

$$I(x) = I(0)e^{-\alpha x} \tag{7.21}$$

where $I(0)$ is the intensity at a point $x = 0$, $I(x)$ the intensity at the position x and α is the absorption coefficient of the medium. The absorption coefficient of air is rather low; some materials show a much larger absorption and can be used to create

In this simple configuration the line is created by a rotating mirror. The projected line follows the shape of the object; this pattern is viewed by a camera. From the known positions of the mirror, the light plane and the camera, the shape of the intersection between the object and the light plane can be reconstructed. In general not all points of the object are illuminated, and not all illuminated points may be viewed by the camera. Those shadows prohibit capturing the whole cross section. When this is not acceptable for the application, a second triangulation system can be used operating under a different angle. Another option is a set of mirrors through which the scanning plane transmitted by a single source is projected from two sides onto the object; the same mirrors can be used to redirect the light from two sides to the common receiver. This solution is proposed in, for instance, Ref. [52], in which the scanning is performed by acousto-optic deflectors. Evidently the two scans should be complementary to build up a complete image of the object. A similar solution with two mirrors is described in Ref. [53]; in this example, elliptical mirrors are used, which allows dynamical reconfiguration of the geometry. This results in a better depth resolution and reducing occlusion effects. Finally we mention a laser range finder that uses a multifaceted pyramidal mirror. This mirror acts as a double scanner, one for the transmitting beam and one for the reflected beam. The system, of which design details can be found in Ref. [54], features high speed, high resolution and reduction of the shadow effect.

Much research is being done to monitor the condition of buried pipes. In Chapter 6 magnetic methods have been discussed, but optical systems are being developed as well. Pipe robots carrying various sensors travel through the pipes under test, collecting important data about the pipe's condition (i.e. obstacles, corrosion, leaks and mechanical defects). The smaller the diameter, the more severe the requirements with respect to size, power efficiency (when battery powered) and costs. As an example of an optical test system for small pipe diameters we refer to Ref. [55]. This work concerns an optical system that measures the inner surface of a 10 mm pipe. The sensing system consists of a laser, a scanning mirror, a motor and a PSD, axially assembled to fit inside the pipe. The system enables the detection of surface defects as small as 0.1 mm, by means of triangulation.

Mechanical scanning is relatively slow and forms the major limitation of the acquisition time, in particular for a 2D scanning mechanism. Line projection using cylindrical optics considerably reduces the acquisition time and is therefore widespread in range finding systems. One solution to speed up scanning of the other dimension is an array of transmitters sending a set of parallel lines to the object under test. The resolution is limited by the number of transmitters, which for practical reasons cannot be very high. This solution is only suitable in cases where no high resolution in the scanning direction is required. An example of an electronic scanning system for laser range finding is presented in Ref. [56]. A limited number of parallel lines is projected onto the object. The cross section of each line has a bell-shaped intensity distribution and the lines overlap partly. When two adjacent lines are transmitted simultaneously, the total intensity of the projection shows a peak at a position that depends on the relative intensities of the individual lines.

7.4.5 Navigation

Movement is change in position; hence, it can be measured by any position sensor, provided such sensor has an adequate range and response time. When the movement is beyond the range of a position sensor, for instance as may occur with moving robots and AGVs, other sensing techniques to capture the position are required. A simple optical navigation method is tracking a marked path. The path may consist simply of a contrasting line on the floor. An LED or other light emitter casts a light beam down to the floor and two detectors on either side pick up the reflected light. The vehicle is controlled in such a way that both detectors receive the same light intensity, that is, when they are equal distance from the reflecting (white) line. When the vehicle starts to deviate from the track, either the left or the right detector receives more light, from which a control signal is derived.

When the position of the robot or vehicle along the path must also be determined, the path can be coded along the track, for instance by a continuous pattern of light and dark fields or particular optical codes [64]. Similarly the movement of a free mobile robot can be tracked by some contrasting 2D pattern on the floor, using sensors detecting dark–light transitions or even an onboard camera viewing the floor [65]. Another possibility to measure speed makes use of the 'natural' irregularities of the floor surface. At two positions of the vehicle (front and rear) the random intensity variations are measured; the speed follows from correlation of the two random signals [38]. Absolute position is then obtained by dead-reckoning techniques, starting from a known position, and possibly updated from time to time with reference positions achieved by special markers or beacons.

Beacons for the determination of absolute position can be either passive or active. Passive beacons are for instance reflectors located at precisely known positions. The transmitters and receivers are located on the vehicle. Active beacons transmit signals that are received by the vehicle. Important design considerations are the directivity (or the scanning area) of the transmitter/receiver, the location of the beacons and the identification of beacons (the vehicle must be able to recognize the active beacons).

As an illustration of optical sensing for navigation, Figure 7.35 shows a robot with a multi-sensor navigation system [66]. This robot has been designed for automatic assembly of industrial and household products. Coarse navigation of this robot is performed by fixed cameras looking to optical markers on the robot. For fine positioning of the robot (to perform proper assembling tasks) a special optical sensing system has been developed using three light-emitting beacons (Figure 7.36).

The sensing system on the robot measures the angles between two beacon pairs: α between B_1 and B_2, and β between B_2 and B_3. Then the robot must be located at a point P lying on the crossings of the two circles c_1 through (B_1,B_2,P) and c_2 through (B_2,B_3,P). From the known beacon positions and the two angles, point P can be calculated. The sensing system on the robot consists of a rotating photo diode (four turns/s), connected to an encoder that provides information on the angle of the optical axis with respect to the robot co-ordinates. During one complete turn, Similarly the relation between dielectric displacement D and electrical field strength E is given by

$$D = \varepsilon \cdot E \quad (\mathrm{C/m^2}) \quad \text{or} \quad E = \left(\frac{1}{\varepsilon}\right) \cdot D \quad (\mathrm{V/m}) \tag{8.11}$$

where ε is the dielectric constant or permittivity (F/m). In piezoelectric materials mechanical and electrical quantities are coupled. A two-port model of a piezoelectric system is shown in Figure 8.3.

The two-port equations of this system are, in general

$$\begin{aligned} S &= f_1(T, D) \\ E &= f_2(T, D) \end{aligned} \tag{8.12}$$

and for linear systems in particular

$$\begin{aligned} S &= s^D \cdot T + g \cdot D \\ E &= -g \cdot T + \left(\frac{1}{\varepsilon^T}\right) \cdot D \end{aligned} \tag{8.13}$$

that is, Eqs (8.10) and (8.11) extended by the parameter g ($\mathrm{m^2/C}$ or V m/N). Equation (8.13) is the *constitutive equation* of the piezoelectric system. Mathematically s^D is the partial derivative of S to T at constant D, and ε^T the partial derivative of D to E at constant mechanical tension T. From a physical point of view s^D is the compliance at open electrical terminals ($\Delta D = 0$), and ε^T is the permittivity at open mechanical terminals ($\Delta T = 0$). The latter means that the material can freely deform: it is not clamped. At open electrical terminals the voltage generated by an applied force is $E = -gT$, which explains the name piezoelectric voltage constant for g.

Further as we have seen before, the impedance of the connecting circuits (both mechanical and electrical) influence the properties of the material and hence the piezoelectric behaviour. For instance the material's stiffness is higher at short-circuited electrical terminals than when these terminals are open.

Another set of constitutive equations is:

$$\begin{aligned} S &= f_3(T, E) \\ D &= f_4(T, E) \end{aligned} \tag{8.14}$$

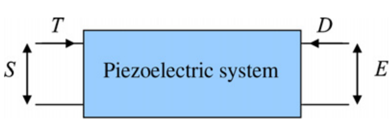

Figure 8.3 Two-port model of a piezoelectric system.

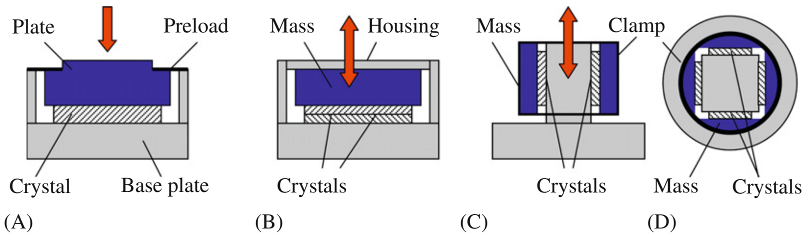

Figure 8.5 Principle of piezoelectric sensors (A) force sensor (longitudinal), (B) compression type accelerometer, (C) shear-type accelerometer, (D) top view of (C); the arrows show the main axis of sensitivity; electrical terminals are not shown.

the active area of the membrane – is mechanically transferred to the crystal. Piezoelectric pressure sensors are also sensitive to acceleration, because the mass of the housing produces an inertial force on the crystal when accelerated. For applications where pressure has to be measured in a vibrating or otherwise moving environment, special pressure sensors are designed with a compensating crystal, to minimize the acceleration sensitivity.

An accelerometer consists basically of one or more piezoelectric crystals and a *proof mass* (or *seismic mass*). Here, also, the mass is fixed by a preload onto the crystal. Figure 8.5B shows a compression mode accelerometer. This example comprises two crystals, mounted back to back between the seismic mass and the base plate. One electrode is connected to the common surface of the crystals, the other to the housing. In this configuration the crystals are electrically in parallel and mechanically in series, resulting in a doubled sensitivity. Moreover some compensation for common interferences is achieved. The shear-type accelerometer shown in Figure 8.5C and D contains four crystals mounted on a rectangular post. The masses and crystals are fixed between a centre post and a clamping ring (as preload).

The absence of moving parts allows a piezoelectric sensor to be mounted in a robust package and hermetically sealed. Figure 8.6 gives an impression of the real construction of the two accelerometer types shown in Figure 8.5.

In both cases the main axis is in the vertical direction, where the sensitivity has a maximum value. Ideally the sensitivity in the orthogonal direction is zero. Commercial sensors have a non-zero sensitivity in directions perpendicular to the main axis, due to construction tolerances (not well-aligned crystals) or the cross sensitivity of the crystal itself. The crystal type and orientation should be chosen such that the highest value of d is in line with the main axis and zero for all other directions. For instance when d_{15} in the shear type is responsible for maximum output (in the *1*-direction), the values of d_{12} and d_{13} should be zero, since these forces normal to the crystal may not generate an electrical output in that direction. Commercial accelerometers have a cross sensitivity that is just a few percent of the main sensitivity.

When two or all three components of a force or acceleration vector have to be measured, multi-axis sensors should be applied. The earlier three-axes versions

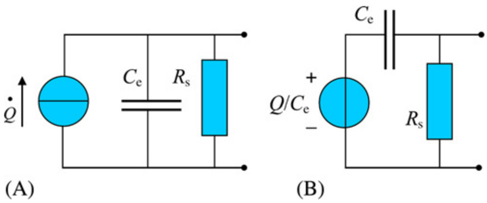

Figure 8.9 Two equivalent models of a piezoelectric sensor: (A) current source and (B) voltage source.

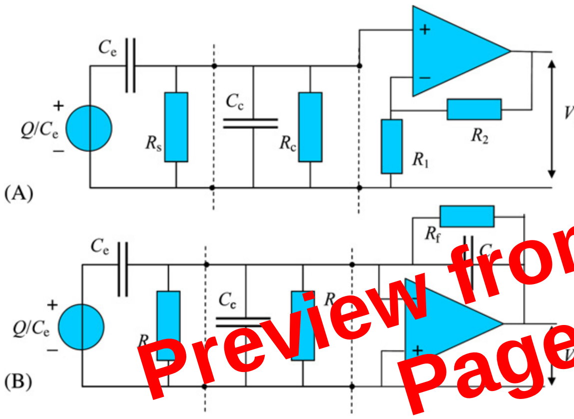

Figure 8.10 Interfacing a piezoelectric sensor: (A) voltage amplifier and (B) charge amplifier.

Figure 8.10 shows examples of these interface circuits, together with models for the sensor and the connecting cable (the part between the dotted lines). It will be shown that the cable capacitance can have a substantial influence on the signal transfer of the system, due to the capacitive character of the sensor. The sensor is modelled by the voltage source model of Figure 8.9B.

We calculate, for both interface circuits, the total signal transfer, using the formula for operational amplifier configurations in Appendix C. First we consider the voltage readout (Figure 8.10A). The transfer of this circuit is given by

$$H_v = \frac{V_o}{Q/C_e} = A \cdot \frac{j\omega R_p C_e}{1 + j\omega R_p (C_e + C_c)} = A \cdot \frac{C_e}{C_e + C_c} \cdot \frac{j\omega R_p (C_e + C_c)}{1 + j\omega R_p (C_e + C_c)} \tag{8.32}$$

where $R_p = R_s // R_c$ (parallel connection) and $A = 1 + R_2/R_1$, the gain of the non-inverting amplifier. Obviously the transfer shows a high-pass character (Figure 8.11A): for frequencies satisfying $\omega R_p(C_e + C_c) \gg 1$ the transfer equals $C_e/(C_e + C_c)$, so it is frequency independent. The cable capacitance C_c causes signal attenuation. Hence the total signal transfer depends on the length of the cable. This requires recalibration each time the connection cable is replaced.

The interface circuit of Figure 8.10B is better, in this respect. Due to the virtual ground of the operational amplifier the voltage across the sensor and the cable is kept at zero: neither the cable impedance nor the input impedance of the amplifier influences the transfer, and hence can be ignored. Assuming an ideal operational the electrical contact surfaces and a proper interfacing. Charge amplifiers are preferred in most cases.

In this section we present some examples of special designs of piezoelectric sensors for various applications.

8.4.1 Stress and Pressure

Axially stressed and poled PVDF shows a rather strong piezoelectric effect (Table 8.2). A straightforward design of a pressure sensor using this property is a sheet of PVDF on top of a rigid backing, making use of lateral deformation. However the flexibility of PVDF allows the material to bend easily, so it can serve as a membrane, resulting in a much higher pressure sensitivity. Two pressure sensors designed according to this concept are presented in Ref. [15]. A circularly shaped sheet of piezoelectric PVDF 5 mm in diameter and about 25 μm thickness, with metallization layers on both sides, is clamped between two parts of a likewise circular housing made up of PVDF. The sensor is applied in pneumatic and hydraulic systems, for pressures up to 200 kPa and temperatures up to 125°C and is resistant to a wide range of chemicals. Many properties of the prototype have been determined experimentally: sensitivity to temperature and humidity, response time (less than 10 μs), frequency range and stability over time (i.e. aging and creep). As outlined in preceding sections, the sensor element is also pyroelectric, so it responds to temperature changes as well, and shows a high-pass characteristic that depends on the temperature-dependent loss factor (resistance) of the membrane.

Ice deposition, for instance on overhead power transmission lines and power pylons, may cause much damage, and it is useful to study the adhesion at different conditions. In Ref. [16] PVDF is used to measure interfacial stress between an ice layer deposited on an aluminium substrate. Clearly the small thickness of PVDF as well as its flexibility makes this sensing material an excellent candidate for this task. It is shown that an embedded PVDF layer provides useful information about ice de-bonding stress and propagation.

Many examples can be found in literature illustrating the versatility of PVDF as a means to measure pressure and obtain pressure images. For instance Ref. [17] describes a system for measuring the dynamic (normal) pressure between a car's tyre and the ground, and in Refs [18,19] the material is used for automatic verification of hand-written signatures.

8.4.2 Acceleration

Most piezoelectric accelerometers for industrial application are designed according to Figure 8.5. The present market offers a vast variety of types for almost any application. However proposals for alternative designs aiming at even better performances or for particular applications appear on a regular basis in scientific literature. For instance the advent of piezoelectric PVDF and the fast development of microtechnology motivate researchers for looking to new solutions for both old and new measurement problems.

where c is the stiffness (modulus of elasticity) and ρ the specific mass of that material. For ideal gases the expression for the speed of sound is:

$$v_a = \sqrt{\frac{c_p}{c_v} \cdot \frac{p}{\rho}} = \sqrt{\frac{c_p R}{c_v M} \Theta} \tag{9.4}$$

where Θ is the absolute temperature, R the gas constant and M the molecular mass. Substitution of numerical values for air yields:

$$v_a = 331.4(1 + 1.83 \cdot 10^{-3} \vartheta) \text{ m/s} \tag{9.5}$$

with ϑ the temperature in °C. Hence, as a rule of thumb, the temperature coefficient of v_a is about 2% per 10 K. The influence of air humidity is relatively small and is only relevant when high accuracy is required.

The speed of sound is roughly 10^6 times lower than the speed of light. Hence ToF is easier to measure, but consequently the measurement time is longer. Moreover the much larger wavelength of sound waves compared to light waves makes it more difficult to manipulate with sound beams (for instance focussing or the creation of a narrow beam).

Acoustic velocity differs from particle velocity u. The latter follows from Euler's equation for the acceleration of a fluid:

$$a = -\frac{1}{\rho} \nabla p \rightarrow \frac{du}{dt} = -\frac{1}{\rho} \frac{dp}{dx} \rightarrow u = -\frac{1}{\rho} \int \frac{dp}{dx} \, dt \tag{9.6}$$

for the particle velocity in the x-direction.

9.1.3 Acoustic Damping

An acoustic wave is attenuated by molecular absorption of sound energy and by scattering: the wave gradually looses energy when propagating. For a plane wave Beer's law applies:

$$P_o = P_i e^{-\alpha \Delta x} \tag{9.7}$$

P_i is the acoustic power at a certain place x in space, P_o is the power at a place Δx farther in the direction of propagation. The attenuation or damping coefficient α comprises two effects: absorption loss and scattering loss. In solids the first effect is proportional to frequency, and the second proportional to the square of the frequency. In gases the second term dominates, hence $\alpha = a \cdot f^2$: the attenuation of the wave increases with squared frequency. Acoustic damping also depends on gas composition. For air it means that acoustic attenuation is affected by the air humidity. Table 9.1 shows some values of damping in dry and humid air.

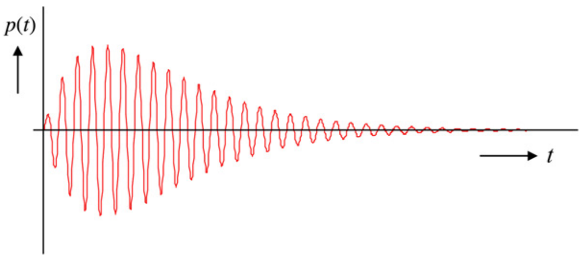

Figure 9.6 Typical response of a PE transducer to an input burst signal.

To overcome the enormous mismatch in airborne applications, the transmission efficiency is improved by placing a precisely shaped piece of an acoustically soft material on the front side of the ceramic element, for instance a tiny horn as shown in Figure 9.5B. A horn or cone also affects the frequency range as well as the radiation field of the transducer [1, p. 102]. Most commercial low-cost piezoelectric transducers are provided with such a horn.

Poled PVDF has a relatively low acoustic impedance (see Table 9.2), making it an attractive material for ultrasound applications. The flexibility of the material allows operation in bending mode, whereas the ceramic sensors usually operate in thickness mode [9,10]. Other piezoelectric materials are being studied for their suitability as acoustic transducer, for instance porous films with artificially electric dipoles [11].

Piezoelectric transducers operate best at resonance: the high stiffness of ceramic materials results in a narrow bandwidth (in other words, a high mechanical Q-factor). The resonance frequency is determined by the dimensions of the crystal and, to a lesser extent, by the matching elements. Popular frequencies are 40 and 200 kHz. Many acoustic detection systems use tone burst signals. A typical response of a piezoelectric transducer to a burst signal at the resonance frequency is shown in Figure 9.6. Apparently the output burst shows strong distortion at the edges, due the narrow bandwidth of the device. This has important consequences for the determination of the ToF measurement, as will be explained in Section 9.3.

9.2.4 Arrays

As shown in Section 9.2.1 the beam width of a sound emitter is determined by its lateral dimensions and the frequency of the emitted sound signal. With a properly designed (linear) array of transducers a much narrower beam can be obtained, using spatial interference of the individual acoustic signals (Figure 9.7).

For points in the hemisphere where the travelled distances differ a multiple of the wavelength, the waves add up (for instance in P). In other points, as in Q, the waves (partly) cancel. Moreover the direction of this beam can be controlled over a limited range ($\pm 30^\circ$) with the phase difference of the applied electrical signals. Similarly the sensitivity angle of an acoustic transducer operating in receiver mode can be narrowed using an array of such receivers. At the receiving side the individual outputs are added after a properly chosen time delay. By controlling the delay

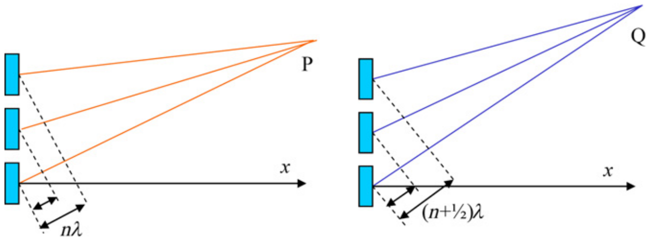

Figure 9.7 Beam narrowing and steering by phased arrays.

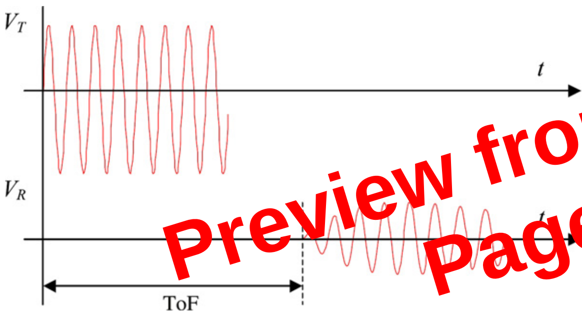

Figure 9.8 Distance determination based on ToF.

times the main sensitivity axis can be varied over a limited angle. The application of this general principle to acoustic transducers has already been described in earlier literature [12–15]. Due to improved technology phased arrays are gaining interest for acoustic ranging and imaging [16–19]. They eliminate the need for a mechanical (hence slow) scanning mechanism but have a limited angular scanning range. In general at an increasing deflection angle the amplitude of the side lobes (Figure 9.3) increases as well, deteriorating the beam quality.

9.3 Measurement Methods

The majority of acoustic distance measurement systems are based on the ToF method. A measurement based on distance-dependent attenuation (as with light) is rather susceptible to environmental influences and the absorption properties of the objects involved. In most acoustic ToF systems the elapsed time between the excitation of a sound pulse and the arrival of its echo is measured (Figure 9.8). Since the transduction effects are reversible, the transducer can be switched from transmitter to receiver mode, within one measurement. The travelled distance x follows directly from the ToF t, using the relation $v_a = x/t$ (for direct travel) or $v_a = 2x/t$ for echo systems (where the sound wave travels twice the distance).

The sound signal can be of any shape. Most popular are the burst (a number of periods of a sine wave), a continuous wave with constant frequency (CW) and an

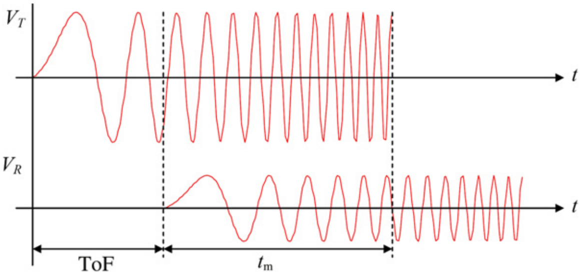

Figure 9.12 FMCW; during t_m the frequency difference is proportional to the ToF.

frequency. The spectrum of this modulated signal contains the components ω_c, $\omega_c + \omega_s$ and $\omega_c - \omega_s$ (Figure 9.11B).

All components should fall within the bandwidth of the transducer, even when using a piezoelectric type. It means that the period time of the 'envelope' is much larger than the wavelength of the carrier. The received wave, delayed over the ToF τ, is written as

$$V_R = A\{(1 + m\cos\omega_s(t-\tau))\cos\omega_c(t-\tau)\} \tag{9.19}$$

The phase difference of the carrier is $\phi_c = \omega_c\tau$; that of the envelope $\phi_s = \omega_s\tau$. So for a ToF corresponding to one period (2π) of the envelope, the carrier has shifted over ω_c/ω_s as many periods. Obviously the unambiguous range is increased by a factor of ω_c/ω_s when using the phase of the envelope signal (to be reconstructed by for instance a rectifying circuit). The method has the advantage of the continuous mode, but with enlarged range. Here, also, the accuracy in the ToF measurement is limited by the resolution of the phase measurement and the uncertainty in sound velocity, according to Eq. (9.1). In Ref. [32] the carrier and modulation frequencies are 40 kHz and 150 Hz, respectively, corresponding to a low-frequency period of about 2 m. Using a digital phase detection circuit, an accuracy of 2 mm over a 1.5 m range was obtained.

9.3.3 Frequency-Modulated Continuous Waves (FMCW)

The unambiguous range of the continuous mode can be increased by frequency modulation of the sound wave. Figure 9.12 shows the transmitted and the received signals (sometimes called a 'chirp', referring to the whistling of some birds).

Assume the frequency varies linearly with time, starting from a lowest value f_L, and increasing with a rate k (Hz/s): $f(t) = f_L + k \cdot t$. Due to the time delay of the reflected wave, its frequency in the same time frame is $f(t) = f_L + k \cdot (t - \tau)$. At any moment within the time period t_m where the signals occur simultaneously, the frequency difference between the transmitted and the received wave equals:

$$\Delta f = f_L + k \cdot t - \{f_L + k(t-\tau)\} = k \cdot \tau \tag{9.20}$$

with τ the ToF. The distance to the reflecting surface follows from

$$x = \frac{1}{2} v_a \tau = \frac{v_a}{2k} \cdot \Delta f \tag{9.21}$$

Apparently this distance is directly proportional to the frequency difference and unambiguous over a time t_m as illustrated in Figure 9.12. The CWFM method combines the advantages of the first two methods: the transmitted signal is continuous (favourable for the S/N ratio, uninterrupted distance information) and the range is larger (determined by the frequency sweep). Disadvantages over the burst method are the more complex interfacing and the fact that only wide-band transducers can be applied since the frequency varies. Further since transmission is essentially continuous, two transducers are required for simultaneous transmitting and receiving.

Important parameters of a CWFM system are the sweep length, the start and stop frequencies and the sweep rate k (in Hz/s). The sweep length should correspond to the largest distance to be measured, to guarantee sufficient time overlap between the transmitted and reflected waves. The start and stop frequencies are set by the characteristics of the transducer. Electrostatic transducers for low-cost applications have a rather limited frequency range, typically from 50 to 100 kHz, so just one octave (a factor of 2). Usually the size of an electrostatic transducer in this frequency range is larger than that of a piezoelectric transducer. This may be a disadvantage from the viewpoint of construction; however, as has been shown in Section 9.2.1, the half angle is much smaller, contributing to a better S/N ratio of the echo signal.

According to Eq. (9.21) the distance measurement requires the determination of the frequency difference between transmitted and received signals. Usually this is done by some correlation procedure, performed in the digital signal domain (see for instance Refs [33,34]). This has the advantage of noise reduction (depending on the type of filtering), but the disadvantage of the need to digitize and sample both signals. However the procedure is essentially based on the multiplication of both signals, followed by some filtering. The basic goniometric relation of Eq. (9.22) is used:

$$\cos\alpha \cdot \cos\beta = \frac{1}{2}\left\{\cos(\alpha+\beta) + \cos(\alpha-\beta)\right\} \tag{9.22}$$

So multiplying two sine waves results in a signal containing sum and difference frequencies. Now let the transmitted FM signal be:

$$x_T(t) = A\cos[2\pi(f_L + k \cdot t)t] \tag{9.23}$$

The reflected wave signal delayed over a time τ becomes:

$$x_R(t) = B\cos[2\pi\{f_L + k(t-\tau)\}(t-\tau)] \tag{9.24}$$

and inspection by touch using ultrasound-based tactile sensors. A series of examples for process monitoring using ultrasound concludes this chapter.

Acoustic Imaging

The goal of any imaging system is the acquisition of an electronic picture from which information about an object's size and shape can be inferred, preferably including as many details as feasible. In an adverse environment in particular, acoustic imaging could be a good alternative for the well-known optical imaging with cameras. Since the spatial resolution of an acoustic distance measurement is relatively low, some method of scanning is required to obtain detailed shape information. The performance of an ultrasonic recognition system is limited by the nature of sound waves, the properties of the transducers, the quality of the scanning system, the number of transducers and the signal processing. In general the acquired image data are sparse and not accurate, which complicates recognition. However in many applications the resolution requirements are not severe.

Scanning can be performed in various ways. Either the objects move relative to fixed transducers, or the transducers are moving relative to a fixed object. In a robotics environment a robot can perform the scanning, thereby making use of the many d.o.f. of the manipulator to perform 3D scanning all around the object [60]. The encoders in the joints provide the orientation of the scanning range finder. A height map is obtained by Cartesian scanning in a plane above the scene. Sideways scanning enables the reconstruction of the circumference of an object, from the position and orientation of the flat or curved sides [61–63]. Several processing algorithms have been proposed to reconstruct the object shape from the obtained data or to classify the objects, including digital filtering, holographic algorithms and artificial neural networks [64–66].

In all these examples mechanical scanning is used. The recognition of objects from a fixed sensory system has been studied as well. Usually such systems comprise an array of transducers with relatively large spacing [67]. For an accurate characterization of large objects (3–10 m), a set of transmitters/receivers is positioned around the object in such a way that together they cover the whole area of interest. An accuracy analysis of such an approach can be found in Ref. [68]. One of the difficulties in such a system is to distinguish between multiple objects and between various shapes returning similar echo patterns. These problems are solved using a set of transmitters and/or receivers at known relative positions and proper signal processing.

For the classification of objects with known shape, the 'signature method' can be applied [62,69,70]. In this approach no attempt is made to reconstruct the object shape from the range data. Instead the system is taught to associate each object class with a characteristic echo pattern. A typical problem to be overcome in this approach is to obtain image data that are position and orientation invariant. On the other hand different but similarly shaped objects may result in echo patterns that are also similar. Correct classification on the basis of the echo patterns then requires special processing techniques, as described, for instance in Ref. [71]. For

[7] G. Caliano, A. Savoia, A. Caronti, V. Foglietti, E. Cianci, M. Pappalardo: Capacitive micromachined ultrasonic transducer with an open-cells structure, Sens. Actuators A, 121 (2005), 382–387.
[8] D. Callens, Chr. Bruneel, J. Assaad: Matching ultrasonic transducer using two matching layers where one of them is glue, NDT&E Int., 37 (2004), 591–596.
[9] F. Harnisch, N. Kroemer, W. Manthey: Ultrasonic transducers with piezoelectric polymer foil, Sens. Actuators A, 25–27 (1991), 549–552.
[10] L. Capineri, M. Calzolai, S. Rocchi, A.S. Fiorillo: A digital system for accurate ranging with airborne PVDF ultrasonic transducers; Proceedings of the ISMCR '95 Conference, Smolenice, Slovakia, 1995, pp. 71–74.
[11] A. Jiménez, Á. Hernández, J. Ureña, M.C. Pérez, F.J. Álvarez, C. De Marziani, J.J. García, J.M. Villadangos: EMFi-based ultrasonic transducer for robotics applications, Sens. Actuators A, 148 (2008), 342–349.
[12] W.L. Beaver: A method of three-dimensional electronic focussing and beam steering using electronic delay lines; Proceedings of Ultrasonics Symposium, 22–24 September 1975, Los Angeles, CA, pp. 88–90.
[13] R.A. Mucci: A comparison of efficient beamforming algorithms, IEEE Trans. Acoust. Speech Signal Process., ASSP–32(3) (1984), 548–558.
[14] J.T. Walker, J.D. Meindl: A digitally controlled CCD dynamically focussed phased array; Proceedings of Ultrasonics Symposium 22–24 September, 1975, Los Angeles, CA, pp. 80–83.
[15] K. Higuchi, K. Suzuki, H. Tanigawa: Ultrasonic phased array transducer for acoustic imaging in air; IEEE Ultrasonic Symposium, 1986, pp. 559–562.
[16] K. Yamashita, L. Chamsomphou, H. Nishimoto, M. Okuyama: A new method of position measurement using ultrasonic array sensor without angular scanning, Sens. Actuators A, 121 (2005), 1–5.
[17] R. Queirós, R.C. Martins, P.S. Girão, A.Cruz Serra: A new method for high resolution ultrasonic ranging in air; Proceedings XVIII IMEKO World Congress, Rio de Janeiro, Brazil, 17–22 September 2006.
[18] K. Yamashita, H. Katata, M. Okuyama: High-directivity array of ultrasonic micro sensor using PZT thin film on Si diaphragm; Transducers 01/Eurosensors XV, Munich, Germany, 10–14 June 2001.
[19] T. Yamaguchi, H. Kashiwagi, H. Harada: A multi-beam ultrasonic echo location system for multi-object environment; ISMCR'99, Tokyo, Japan, 10–11 June 1999; in: Proceedings of IMEKO-XV World Congress, Osaka, Japan, 13–18 June 1999, Vol. X, pp. 1–6.
[20] M. Parrella, J.J. Anaya, C. Fritsch: Digital signal processing techniques for high accuracy ultrasonic range measurements, IEEE Trans. Instrum. Meas., 40(4) (1991), 759–763.
[21] D. Marioli, C. Narduzzi, C. Offelli, D. Petri, E. Sardini, A. Taroni: Digital time-of-flight measurement for ultrasonic sensors, IEEE Trans. Instrum. Meas., 41(1) (1992), 93–97.
[22] J.M. Martín Abreu, R. Ceres, T. Freire: Ultrasonic ranging: envelope analysis gives improved accuracy, Sensor Rev., 12(1) (1992), 17–21.
[23] E. van Ginneken, P.P.L. Regtien: Accurate ultrasonic distance measurement using time-of-flight; AIM '92 (Australasian Instrumentation and Measurement Conference), Auckland, New Zealand, 24–27 November 1992, pp. 377–380.
[24] F.E. Gueuning, M. Varlan, C.E. Eugène, P. Dupuis: Accurate distance measurement by an autonomous ultrasonic system combining time-of-flight and phase-shift methods, IEEE Trans. Instrum. Meas., 46(6) (1997), 1236–1240.

Appendix A

Symbols and Notations

Sensors operate at the boundary of two physical domains. Despite international normalization of symbols for quantities and material properties, notations for physical quantities are not unambiguous when considering the various disciplines, which each have their own system of notation. This appendix offers a brief review of quantities in the electrical, thermal, mechanical and optical domains, together with their symbols and units, as used in this book. Further, where useful, relations between quantities are listed as well.

A.1 The Electrical Domain

A sensor produces an electrical output. Table A.1 displays the most important energetic quantities used in the electrical domain. The table also shows magnetic quantities, since they are closely connected to electrical quantities. Relations between these variables are discussed in Chapter 6.

Table A.2 shows the major properties for the electrical domain.

Conductivity is the inverse of resistivity, and conductance the inverse of resistance. The electric permittivity ε is the product of the permittivity of free space (vacuum) ε_0 and the relative permittivity (or dielectric constant) ε_r. Similarly the magnetic permeability μ is the product of the magnetic permeability of vacuum μ_0 and the relative permeability μ_r. Numerical values of ε_0 and μ_0 are:

$$\varepsilon_0 = (8.85416 \pm 0.00003) \cdot 10^{-12} (\mathrm{F/m})$$
$$\mu_0 = 4\pi \cdot 10^{-7} (\mathrm{V\ s/A\ m})$$

The relative permittivity and permeability account for the dielectric and magnetic properties of a material.

In the electrical domain, some particular variables apply, associated to properties of (electrical) time-varying signals. They are listed in Table A.3.

The duty cycle is defined as the high–low ratio of one period in a periodic pulse signal. It varies from 0% (whole period low) to 100% (whole period high). A duty cycle of 50% refers to a symmetric square wave signal.

Table A.1 Electrical and Magnetic Quantities

Quantity	Symbol	Unit
Electric current	I	A (ampère)
Current density	J	A m^{-2}
Electric charge	Q	C (coulomb) = A s
Dielectric displacement	D	C m^{-2}
Electric field strength	E	V m^{-1}
Potential difference	V	W A^{-1}
Magnetic flux	Φ	Wb (weber) = J A^{-1}
Magnetic induction	B	T (tesla) = Wb m^{-2}
Magnetic field strength	H	A m^{-1}

Table A.2 Electrical Properties

Electrical Property	Symbol	Unit
Resistivity	ρ	Ω m
Resistance	R	Ω (ohm)
Conductivity	σ	S m^{-1} = Ω^{-1} m^{-1}
Conductance	G	S (siemens)
Capacitance	C	F (farad)
Self-inductance	L	H (henry)
Mutual inductance	M	H
Permeability	$\mu = \mu_0 \cdot \mu_r$	H m^{-1}
Relative permeability	μ_r	–
Permittivity	$\varepsilon = \varepsilon_0 \cdot \varepsilon_r$	F m^{-1}
Dielectric constant	ε_r	–

Table A.3 Signal Variables

Signal Quantity	Symbol	Unit
Time	t	s (second)
Frequency	f	Hz (hertz) = s^{-1}
Period	T	s
Phase difference	ϕ	rad (radian)
Duty cycle	δ	–
Pulse width	τ	s

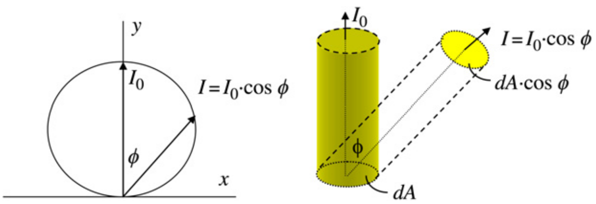

Figure A.3 Cosine law of Lambert.

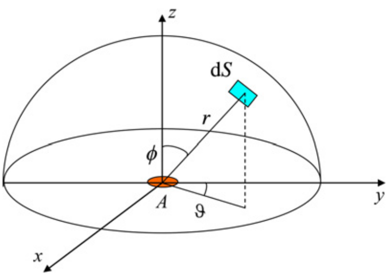

Figure A.4 Derivation of the total radiant energy from a unit surface with Lambertian emission.

A.4.2 Radiant Energy from a Unit Surface with Lambertian Emission

If a surface is a perfect diffuser (its radiance is not a function of angle), then $I = I_0 \cdot \cos \phi$ (Figure A.3).

This is known as the cosine law of Lambert. Many surfaces scatter the incident light in all directions equally. A surface that satisfies Lambert's law is called Lambertian.

All radiant energy emitted from a surface A with radiant emittance E_s has to pass a hemisphere around that surface (Figure A.4).

The area of surface element dS on the hemisphere, corresponding with a solid angle dΩ equals:

$$\mathrm{d}S = r^2 \mathrm{d}\Omega = r^2 \sin \phi \mathrm{d}\phi \mathrm{d}\vartheta \tag{A.9}$$

where r is the radius of the hemisphere. According to its definition in Table A.10 the radiant emittance from the emitting surface A in the direction of dS is:

$$\frac{\mathrm{d}P_s}{\mathrm{d}A} = L \cos \phi \mathrm{d}\Omega \tag{A.10}$$

Since all radiant energy from A passes the hemisphere, the total radiant emittance of A equals the integral of dP_s/dA over the full surface area of the hemisphere. Further since for a Lambertian surface the radiant emittance is independent of the direction, L is constant. So

Since we consider only small variations, the variables S, D, B, and $\Delta\sigma$ are approximated by linear functions, so

$$\begin{bmatrix} \mathrm{d}S(T,E,H,\Theta) \\ \mathrm{d}D(T,E,H,\Theta) \\ \mathrm{d}B(T,E,H,\Theta) \\ \mathrm{d}\sigma(T,E,H,\Theta) \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial S}{\partial T}\right)_{E,H,\Theta} & \left(\frac{\partial S}{\partial E}\right)_{T,H,\Theta} & \left(\frac{\partial S}{\partial H}\right)_{T,E,\Theta} & \left(\frac{\partial S}{\partial \Theta}\right)_{T,E,H} \\ \left(\frac{\partial D}{\partial T}\right)_{E,H,\Theta} & \left(\frac{\partial D}{\partial E}\right)_{T,H,\Theta} & \left(\frac{\partial D}{\partial H}\right)_{T,E,\Theta} & \left(\frac{\partial D}{\partial \Theta}\right)_{T,E,H} \\ \left(\frac{\partial B}{\partial T}\right)_{E,H,\Theta} & \left(\frac{\partial B}{\partial E}\right)_{T,H,\Theta} & \left(\frac{\partial B}{\partial H}\right)_{T,E,\Theta} & \left(\frac{\partial B}{\partial \Theta}\right)_{T,E,H} \\ \left(\frac{\partial \sigma}{\partial T}\right)_{E,H,\Theta} & \left(\frac{\partial \sigma}{\partial E}\right)_{T,H,\Theta} & \left(\frac{\partial \sigma}{\partial H}\right)_{T,E,\Theta} & \left(\frac{\partial \sigma}{\partial \Theta}\right)_{T,E,H} \end{bmatrix} \cdot \begin{bmatrix} \mathrm{d}T \\ \mathrm{d}E \\ \mathrm{d}H \\ \mathrm{d}\Theta \end{bmatrix} \tag{B.19}$$

Combining Eqs (B.18) and (B.19) results in

$$\begin{bmatrix} \mathrm{d}S \\ \mathrm{d}D \\ \mathrm{d}B \\ \mathrm{d}\sigma \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial^2 G}{\partial T^2}\right)_{E,H,\Theta} & \left(\frac{\partial^2 G}{\partial T \partial E}\right)_{H,\Theta} & \left(\frac{\partial^2 G}{\partial T \partial H}\right)_{E,\Theta} & \left(\frac{\partial^2 G}{\partial T \partial \Theta}\right)_{E,H} \\ \left(\frac{\partial^2 G}{\partial E \partial T}\right)_{H,\Theta} & \left(\frac{\partial^2 G}{\partial^2 E}\right)_{T,H,\Theta} & \left(\frac{\partial^2 G}{\partial E \partial H}\right)_{T,\Theta} & \left(\frac{\partial^2 G}{\partial E \partial \Theta}\right)_{T,H} \\ \left(\frac{\partial^2 G}{\partial H \partial T}\right)_{E,\Theta} & \left(\frac{\partial^2 G}{\partial H \partial E}\right)_{T,\Theta} & \left(\frac{\partial^2 G}{\partial^2 H}\right)_{T,E,\Theta} & \left(\frac{\partial^2 G}{\partial H \partial \Theta}\right)_{T,E} \\ \left(\frac{\partial^2 G}{\partial \Theta \partial T}\right)_{E,H} & \left(\frac{\partial^2 G}{\partial \Theta \partial E}\right)_{T,H} & \left(\frac{\partial^2 G}{\partial \Theta \partial H}\right)_{T,E} & \left(\frac{\partial^2 G}{\partial^2 \Theta}\right)_{T,E,H} \end{bmatrix} \cdot \begin{bmatrix} \mathrm{d}T \\ \mathrm{d}E \\ \mathrm{d}H \\ \mathrm{d}\Theta \end{bmatrix} \tag{B.20}$$

The second-order derivatives in the diagonal represent properties in the respective domains: mechanical, electrical, magnetic and thermal. All other derivatives represent cross effects. These derivatives are pair-wise equal, since the order of

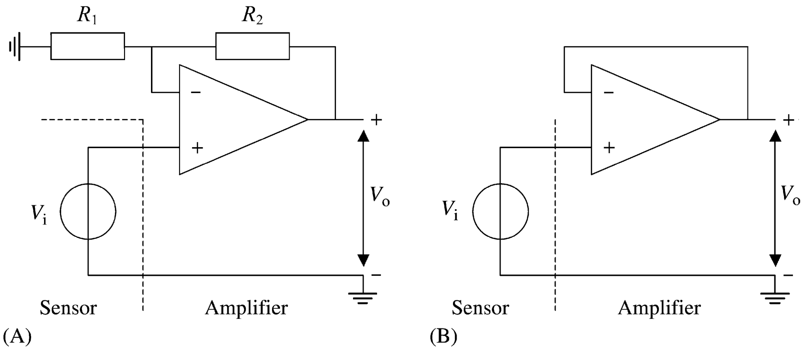

Figure C.3 (A) Non-inverting amplifier, (B) buffer.

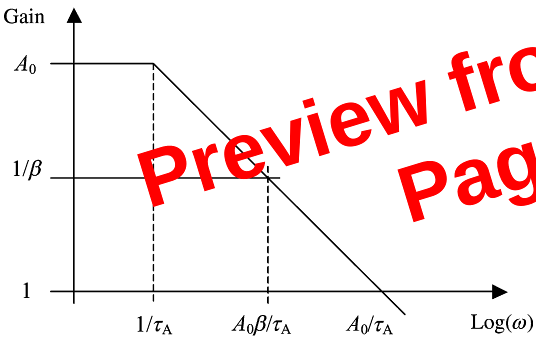

Figure C.4 Frequency characteristic of a non-inverting amplifier: asymptotic approximation according to Eq. (C.9).

$$V_o = \frac{V_i}{\beta}\frac{A_0\beta}{1+A_0\beta} \approx \frac{V_i}{\beta}\left(1-\frac{1}{A_0\beta}\right) \tag{C.8}$$

The scale error is the inverse of the loop gain $A_0\beta$. Taking into account the frequency-dependent gain as in Eq. (C.4), the transfer of the non-inverting amplifier is:

$$V_o = \frac{V_i}{\beta}\frac{1}{1+j\omega\tau_A/A_0\beta} \tag{C.9}$$

Obviously the bandwidth is proportional to $A_0\beta$, and the transfer is inversely proportional to $A_0\beta$, hence the product of gain and bandwidth has a fixed value: the gain-bandwidth product. This relationship is shown in Figure C.4A.

Numerical example. $A_0 = 10^4$, $\beta = 0.01$, so $A_0\beta = 100$; unity gain bandwidth $A_0/\tau_A = 10^6$ rad/s, so the operational amplifier's cut-off frequency is $1/\tau_A = 100$ rad/s and the non-inverting amplifier bandwidth is 10^4 rad/s.

Figure C.5 shows a simulation[1] of the transfer characteristic using a general purpose, low-cost operational amplifier for three gain factors: 1 (buffer), 100 and

[1] Type μA741 (original design by Fairchild, 1968); simulated by PSpice (OrCAD Inc.).

The contribution of the bias current I_b^- can be reduced by inserting a resistance R_3 in series with the non-inverting input terminal, with a value equal to $R_1//R_2$, that is the parallel combination of R_1 and R_2. With this bias current compensation, the term I_b^- in Eq. (C.11) is replaced by I_{off}.

The transfer including multiplicative errors due to a finite gain is:

$$\frac{V_o}{V_i} = -\frac{R_2}{R_1}\frac{A_0\beta}{1 + A_0\beta} \approx -\frac{R_2}{R_1}\left(1 - \frac{1}{A_0\beta}\right) \tag{C.12}$$

which, again, introduces a scale error equal to the inverse of the loop gain. The frequency-dependent gain is given by

$$\frac{V_o}{V_i} = -\frac{R_2}{R_1}\frac{1}{1 + j\omega\tau_A/A_0\beta} \tag{C.13}$$

Here, too, the product of gain and bandwidth is constant. The frequency characteristics are similar to those in Figure C.4. The input resistance becomes:

$$R_{in} = R_1\left(1 + \frac{R_2}{A_0 R_i}\right) \tag{C.14}$$

a value that is almost equal to R_1. Note that this value can be rather low, resulting in an unfavourable load on the transducer.

C.5 Comparator and Schmitttrigger

C.5.1 Comparator

A voltage comparator (or short comparator) responds to a change in the polarity of an applied differential voltage. The circuit has two inputs and one output (Figure C.7A). The output has just two levels: high or low, all depending on the polarity of the voltage between the input terminals. The comparator is frequently used to determine the polarity in relation to a reference voltage.

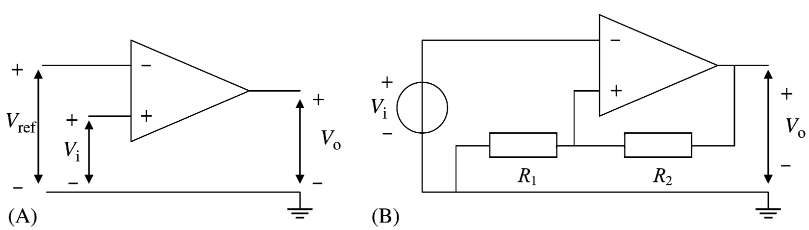

Figure C.7 (A) Comparator, (B) Schmitttrigger.

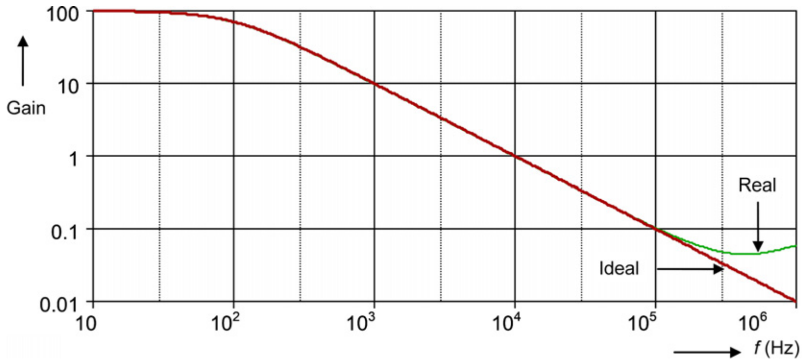

Figure C.10 Frequency characteristic of the modified integrator according to Eq. (C.18) (ideal) and according to a simulation with a real amplifier (real).

For $A = A_0$ and $A_0 \gg 1$ this is:

$$\frac{V_o}{V_i} = -j\omega RC \cdot \frac{A_0}{1 + A_0 + j\omega RC} = -A_0 \frac{j\omega RC/A_0}{1 + j\omega RC/A_0} \tag{C.20}$$

This corresponds to a first-order high-pass characteristic with cut-off frequency at $\omega = A_0/RC$. The differentiation range is limited to this frequency. For much lower frequencies the transfer function is just $-j\omega RC$ and corresponds to that of an ideal differentiator.

However when the frequency dependence of the operational amplifier gain according to Eq. (C.4) is taken into account, the transfer function becomes:

$$\frac{V_o}{V_i} = -j\omega RC \cdot \frac{A_0}{A_0 + j\omega(\tau_A + RC) - \omega^2 \tau_A RC} \tag{C.21}$$

This characteristic shows a sharp peak at $\omega^2 \tau_A \tau / A_0 \approx 1$, for which value the transfer is about $A_0 RC$.

Numerical example. The required gain is 0.1 at 1 rad/s. So the component values R and C should satisfy $1/RC = 10$ rad/s. Suppose the amplifier properties are: $A_0 = 10^5$, unity gain bandwidth 10^5 rad/s. Figure C.11 presents the calculated characteristic for this example. The transfer shows a peak of 10^4 at the frequency 10^3 rad/s.

Figure C.12 shows a simulation for a general purpose operational amplifier with frequency compensation (i.e. it behaves almost as a first-order system). Component values in this simulation are $R = 100$ kΩ, $C = 1$ μF. Differences compared to Figure C.11 are due to different amplifier parameters A_0 and τ_A of the real amplifier.

If the operational amplifier has a second-order cut-off frequency (and most do have), instability may easily occur. Therefore the differentiating range should be limited to some upper frequency, well below the cut-off frequency of the