4 Sensors

4.1 Introduction

This chapter explains how a measurand can modify some property of a sensor in order to produce an electrical signal on which the measurand information is “impressed”. In this way the information is passed to the electrical domain which permits managing it with all the tools available for electrical signal acquisition, transmission, processing and storing.

The description of some sensors will need information already presented in Chapter 3, so that several references to the previous chapter will be done to help those who skipped the reviewing chapter.

Before describing the sensors in detail some elementary electric circuits will be introduced in order to visualize how sensors are included into the circuits which are essential parts of the instruments. Although generic, these ideas are probably new for most readers, and because of this they are introduced early in the chapter to facilitate the understanding of the following material.

The concepts on electronics introduced in Section (4.2) could not only be considered as a simple summary for those with little experience in electric circuits, but they could also be of help when some simple laboratory measurements are required. Among these concepts are the dispositions of sensors in series or parallel according to their impedances, the Wheatstone bridge and the four wire technique, which are classical ways of connecting sensors. Also, a summary on operational amplifiers is presented to demystify these very useful devices for amplifying sensor signals. A simple to build amplifier will be described for students who could require amplifying small signals and do not have any experience in electronics.

The descriptions are first centered on how the elementary electrical devices such as impedances are modified by external parameters (measurands) in order to become a sensor. It would not be possible to understand how sensors work without a clear understanding on how an electrical signal “represents” a non-electrical measurand. We could call these elementary sensors proto sensors because they are the basis or origin of more sophisticated sensors. Grasping these simple devices prepares the readers for understanding any other sensor that they could find in the future.

Basically, any sensor with electrical output can be assimilated to one of these two cases:
1. An electrical generator, as is the case of transducers that transform some type of energy into electrical energy.
2. A variable impedance (resistance, capacitance or inductance).

The measurand, which is the variable that we want to know, is “transformed” into an electrical variable. In the transducers case the measurand produces some kind of
energy that the transducer transforms into electrical energy. In the second case, the
sensor is a variable impedance and the measurand just modifies this impedance.

In the transducer case the first electronic conditioning step is signal amplification
to bring the signal to adequate levels for further treatment. In the second case, the
sensor (variable impedance) must be part of an electric circuit to copy the measurand
information into electrical information.

Therefore, the descriptions of the functioning principle of sensors to be carried out
in this chapter are not merely explanations of sensors used in Environmental Sciences.
They are also intended as a didactic tool to prepare the readers to understand the
working principle of any sensor that they could meet over the course of their careers.

After grasping the simplest ideas on how sensors convert measurand information
into electrical information, more complex sensors are presented which will be the
first stage of the instruments described in the following chapters.

An introduction to oscillators is also presented here. It is essential to understand
how they work because they are the base of clocks, alternating circuits, acoustic
and electromagnetic waves generation and excitation sources for some sensors. The
treatment of this topic has been postponed a bit since some knowledge on sensors
and actuators is needed to make this subject more understandable.

4.2 Introduction to Electric and Electronic Circuits

In order to obtain an electrical signal that can be useful later as part of a measuring
system, sensors must be part of an electric circuit. The most adequate circuit for each
sensor depends on the sensor characteristics. If the sensor generates its own electrical
energy, as it happens with transducers, the first electrical stage is usually a simple
amplifier. If the sensor does not produce any electrical energy it must be included in
an electric circuit whose electrical characteristics can be modified by the measurand.
In this case the sensor’s impedance is modified according to the phenomenon we want
to measure. Then the circuits accompanying sensors have to be selected according to
their electrical characteristics.

Some schematics of circuits are shown below with the purpose of introducing
the reader to the subject. Because they are easier to understand, DC circuits using
resistive sensors are used. In this case sensors change their resistances due to a
change in the measurand. The concepts so introduced will be later extended to AC
circuits where sensors change their impedances due to changes in the measurand.

Figure 4.1a shows a constant voltage battery connected in series with a resistive
sensor and an ammeter which measures current. This kind of circuit is used when the
sensor has a relatively high resistance compared with that of the measuring circuit
(ammeter), thus almost all resistance in the circuit is due to the sensor. Therefore, the
circuit results predominantly sensitive to sensor variations because the resistance of
the measuring circuit can be disregarded (Section (3.7.6)). Changes in the measurand
produce a considerable current modification in agreement with Ohm’s law. The arrow crossing the resistor symbol indicates that it is a resistor that varies according to the external measurand. In this circuit the “exciting” source (which is fixed) is the voltage and the “output” variable signal is the current, on which the measurand information is “copied”.

Another kind of circuit is used when the sensor has a relatively low resistance compared with that of the measuring circuit. The circuit consists of a constant current source (as the exciting source) connected in series with the sensor. The voltage on the sensor (output) is measured by a voltmeter placed in parallel with the sensor (Fig. 4.1b). The voltmeter is assumed to have a high input resistance and practically no current is derived through it. Thus the constant current passing through the variable sensors’ resistance (which is modified by the measurand) produces a variable voltage. Therefore, in this case, the measurand information is “copied” on the voltage (Section (3.7.6)).

**Fig. 4.1:** (a) The sensor resistance is much larger than the ammeter resistance, thus almost all the measured current is due to the sensor variations. (b) The voltmeter resistance is much larger than the sensor resistance then the measured voltage is due to sensor variations.

### 4.2.1 Wheatstone Bridge

The Wheatstone bridge is a high sensitivity circuit often employed to process sensor signals analogically before entering the amplifying stage. It was originally a laboratory circuit composed of four resistors (some of them being the sensing elements), a voltage supply and a galvanometer used as a null-balance meter (Fig. 4.2a).
When the voltage drop between A and B is equal to zero, the galvanometer will indicate zero, and the bridge is said to be “balanced”. The balance condition is achieved when the resistors are related by

\[ \frac{R_1}{R_2} = \frac{R_3}{R_4} \]  

(4.1)

The Wheatstone bridge allows the value of an unknown resistor to be determined when the other three are known and the bridge is led to the balance condition. With the purpose of obtaining the zero current condition through the galvanometer, a set of variable resistors with accurately known resistances should be available. For example (Fig. 4.2b), if \( R_x \) is unknown, \( R_3/R_4 = K \) is a known constant, and \( R_v \) is a variable precise resistor which is adjusted to balance the bridge, then \( R_x \) is given by

\[ R_x = K R_v \]  

(4.2)

A modified version of the bridge is used in modern instruments where the galvanometer is replaced by the input of an instrumentation amplifier. The bridge is not led to balance but a small imbalance is admitted. Also, this bridge has been generalized to measure impedances. In this case the battery is replaced by an AC voltage of adequate known frequency. With this extended version of the bridge, capacitances and inductances can be measured. Both versions of the bridge will be seen later on as part of measuring instruments.

**Fig. 4.2:** (a) Wheatstone bridge, (b) The resistor \( R_v \) is used to measure \( R_x \).

### 4.2.2 Four Wire Technique

Other electrical circuit frequently used with low-resistance resistive sensors is known as *four wire sensing*. Sometimes the sensor resistance is so low that it is comparable to the cable resistance. This can happen when long cables are needed because the sensor is far from the power supply (\( V_s \)) and the measuring instruments. In these
cases the voltage drops through the cables add to the sensor’s voltage and could result a serious problem if the cable resistance vary with temperature, because cable variations become indistinguishable from sensor variations (Fig. 4.3).

The four wires sensing method separates the current supply circuit from the voltage measuring circuit (Fig. 4.4). At first glance it seems amazing that by adding more cables as \( w_1 \) and \( w_2 \), the influence of the cable drops decrease. It happens that the voltmeter is an instrument with high internal resistance and needs a very low current (\( I_m \)). Then, the current in cables \( w_1 \) and \( w_2 \) are very low and so are the voltage drops in series with the voltmeter, thus the voltmeter measures practically only the voltage on the sensor.

**Fig. 4.3:** An ammeter A and a voltmeter V are used to measure the sensor resistance. The cable resistance appears in series with the sensor, thus cable variations are indistinguishable from sensor variations, which are the signal we want to measure.

**Fig. 4.4:** The cables \( w_1 \) and \( w_2 \) separate the current supply (\( I_e \)) circuit from the voltmeter circuit. Because the voltmeter has a high internal resistance, the current \( I_m \) is much lower than \( I_e \), thus proportionally decreasing the voltage drops in series with the voltmeter. Therefore, the voltmeter measures approximately the resistive sensor voltage.
All the concepts explained above relative to circuits with resistances can be extended to cases where alternating voltages and impedances are used.

4.2.3 Operational Amplifier

Operational Amplifiers (OA) are devices very often used in electronics and they will be introduced conceptually to facilitate the understanding of signal amplification in instruments. In general, sensors require some degree of amplification and it would be interesting that researchers could have the ability to employ OA to adapt sensor’s signals to data loggers or other measuring instruments such as testers and oscilloscopes. In order to adequately design with OA it is necessary to have some background on electronics. Nevertheless, a practical circuit that students could use in their projects will be introduced.

Two OA are depicted in Figure 4.5; the triangle represents the OA and the word GAIN indicates the amount of times that the input signal is amplified by the OA. The amplifier has two connections for power supply (+PS and –PS), two inputs (INV and NON-INV) and one output. Also shown are a voltage input signal $V$ and an electrical potential adopted as reference (REF).

This is an OA which requires dual power supply, one positive (+PS) and one negative (–PS), with respect to the REF. The input signal might be the signal from a sensor and could be positive or negative with respect to REF. Because in both cases one input terminal is connected to the REF it is said that the OA has single-end input, in opposition to differential input OA which will be described below.

When the input signal to be amplified ($V$) is introduced between the REF and the NON-INV input, as in Figure 4.5 (a), the output will have an amplitude equal to that of the input signal times the GAIN, and will show the same polarity than the input. This means that if the input is increasing with time, the output will also be increasing.

When the input $V$ is connected to INV the amplitude of the output will also be the input signal times the GAIN, but with the opposite polarity than the input, as shown in Figure 4.5 (b). In other words, if $V$ is positive with respect to REF, the output will be negative.

The interesting capability of an OA is that the gain may be very large ($10^6$) so that very small signals can be greatly amplified. Frequently, most sensors have very low analog signals and OA are essential to amplify sensor signals before converting them into digital ones (see Section (3.6.7), Application example, data logger).

It should be taken into account that the amplifier also amplifies the noise. Then, certain precautions must be taken to avoid noise entering the amplifier.

The NON-INV input has very high input impedance; this means that it will practically take no current from the input signal $V$. This implies that the introduction of the OA will not perturb the signal source. If $V$ is the output of a sensor with high internal impedance (such as a PH meter electrode) we would like to connect it to the
NON-IN V input to disturb the sensor signal as little as possible. Remember that the input of the amplifier is in parallel with the sensor. In contrast, the INV input usually needs some current from V and should be avoided with sensors with high internal impedances.

In sum, the OA is a device which can multiply the input voltage by the GAIN of the amplifier. The output waveform can have the same or opposite polarity than the input, depending if the input signal is connected to the NON-IN V or INV input respectively. The OA requires a power supply for working (two in this case of a dual-power supply OA). The GAIN of the schematics of Figure 4.5 is fixed by external resistors not drawn.

There are other amplifiers in which the input signal is connected between two input terminals (Fig. 4.6); they are called differential OA. They can have two kinds of outputs. One, with only one output connector; the output signal is thus referred to a reference voltage as in Figure 4.5. The other with two output connectors; the signal is obtained between them. In the second case, where inputs and outputs have two terminals, OA are called fully-differential. In fully-differential OA the amplitude of the output voltage (between the two outputs terminals) is the signal at the input terminals times the gain of the amplifier.

For example, a differential input amplifier is usually employed to amplify the signal of a Wheatstone bridge. The input of the amplifier is connected replacing the galvanometer in Figure 4.2b.
Another application could be to replace the ammeter in Figure 4.1a by a low resistance ($R$) and a fully-differential amplifier (Fig. 4.7). Thus, by Ohms’ Law, the input signal to the amplifier is the current times the value of the low resistance (voltage drop across $R$). The output voltage ($V_{\text{out}}$) of the OA is the input voltage times the gain ($G$) of the amplifier (Eq. (4.3)). Therefore,

$$V_{\text{out}} = I R G.$$  \hspace{1cm} (4.3)

In this case, because $R$ is low, the voltage drop across it will be also low, thus $G$ has to be high to obtain a good output signal.

In a similar way, in a real instrument using the circuit of Figure 4.1b, the fully-differential amplifier could replace the voltmeter. In this case perhaps the gain could be low because the voltage on the sensor is large.
An OA extensively used in instruments, which due to its good performance is still recommended for new designs, is the INA 111. This instrumentation amplifier will be presented with a double purpose: to illustrate the previous concepts, and to provide an easy to build amplifier for students which require amplifying small signals and do not have any experience in electronics.

The circuit of Figure 4.8 represents the INA 111 with some external components (http://www.ti.com/lit/ds/symlink/ina111.pdf.). Note that this is an ideal representation because the polarization resistances needed at the input of the amplifier are not shown. They will be added in Figure 4.9. This OA has differential input because it has two inputs ($V_{IN-}$ and $V_{IN+}$) where the input signal (sensor signal) is connected.

The output voltage is obtained on the LOAD, between two terminals: the amplifier’s output ($V_{OUTPUT}$) and the reference (REF); thus, because it has only one output connector it is not a fully differential amplifier. The load may be simply a resistor, the input of a data logger, a voltmeter, an oscilloscope, etc. In this case, the reference is connected to ground, which is the point where the positive and negative voltage supplies +PS and –PS are connected too. Two 0.1 μF capacitors ($C$) are connected in parallel to the power supplies just to reduce the electrical induced noise.

![Fig. 4.8: Schematic general diagram of an INA 111.](image-url)

The GAIN is user selectable; it is the factor by which it is desired to multiply the input voltage to obtain the required output voltage. In the INA 111, this factor is selected
by placing a resistor $R_g$ between two pins of the OA. The output voltage as a function of the input voltage between $V_{IN-}$ and $V_{IN+}$ and the gain ($G$) is shown in Eq. (4.4), where the gain has subsequently been replaced by a function of the resistor’s value $R_g$ expressed in ohms. This equation is provided by the OA manufacturer.

$$V_{OUTPUT} = GAIN \times (V_{IN+} - V_{IN-}) = \left(1 + \frac{50,000}{R_g}\right) \times (V_{IN+} - V_{IN-})$$  \hspace{1cm} (4.4)

Approximate values of $R_g$, using standard resistors are shown in Table 4.1 to obtain different gains (http://www.ti.com/lit/ds/symlink/ina111.pdf.).

**Table 4.1: Gains and standard resistors $R_g$**

<table>
<thead>
<tr>
<th>Desired gain</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_g$ Standard resistor (kΩ)</td>
<td>12.4</td>
<td>5.62</td>
<td>2.61</td>
<td>1.02</td>
<td>0.511</td>
</tr>
</tbody>
</table>

The circuit of Figure 4.8 needs only a few more resistors for practical use. The value and position of these resistors depend on the signal source (sensor) characteristics. For the sake of clarity these resistors are drawn in Figure 4.9, where Figure 4.8 has been simplified (+PS, -PS, $R_g$ and both C are not shown). The two circuits presented could cover many amplification applications. Circuit (a) has a signal source with a large internal resistor $R_s$ such as crystal or ceramic transducers, and requires two large resistors for polarization, one at each input terminal; $R_p = 1,000,000 \, \Omega$ being an appropriate value for both resistors.

![Fig. 4.9: Schematic general diagram for two application cases of an INA 111. In (a) $R_s$ is large and two $R_p$ are needed. In (b) $R_s$ is small and only one $R_p$ is required. For selecting $R_s$ values see the text.](image-url)
Circuit (b) is used with signal sources with small internal resistors $R_s$ such as thermocouples; in this case only one $R_p = 10,000 \, \Omega$ is adequate. Therefore, adding one or two resistors to the circuit of Figure 4.8 a practical amplifier can be built for low or high internal resistance signal sources respectively.

For applications where these circuits do not work appropriately some help could be found in http://www.ti.com/lit/ds/symlink/ina111.pdf.

### 4.3 Resistive Sensors

Resistive sensors are devices that modify their resistance due to changes in a given physical or chemical quantity (measurand) acting on them. As shown in Eq. (3.12) (repeated in Eq. (4.5)) the electrical resistance of a practical resistor depends on the conductivity of the material ($\sigma$), its length ($l$) and cross section ($A$) ($\rho = 1/\sigma$ is the resistivity of the material). Therefore, in order to have a resistive sensor some of these parameters have to vary in response to an external action, so that changes in resistance might be induced by the measurand,

$$R = \frac{\rho l}{A} \quad (4.5)$$

Some examples of resistive sensors are presented below. Sometimes, resistors are built of a material which has a constant resistivity, and mechanical strains induced by the measurand alter their geometry (length or cross section). In this way $R$ varies due to a dimensional change of the resistor (Eq. (4.5)); this is the case of sensors for measuring mechanical stresses, called strain gauges (described below in Section (4.6)).

A different case is when the resistor’s geometry (length and cross section) is kept constant, but the resistivity of the material that forms the resistor changes with a certain external parameter (measurand). For example, if the material forming the resistor were porous and capable of absorbing moisture, $\rho$ in Eq. (4.5) could change because of the water entering the pores of the resistor’s material. Thus, the resistor would change with the amount of water absorbed and a resistive humidity sensor could be obtained (described below in Section (4.9)).

#### 4.3.1 Sensors Based on Potentiometers

So far we have been describing resistors that have two terminals. The potentiometer is a special case of a resistor which has three terminals and its resistance can be varied; it was originally known as rheostat. A schematic of a potentiometer is depicted in Figure 4.10. It consists of a resistive strip with two fixed terminal at the ends and a third terminal (called wiper) that can slide along the strip.
Let us define a number \( n \), \( 0 \leq n \leq 1 \), such that \( n = 0 \) when the sliding terminal touches the upper terminal and \( n = 1 \) when it touches the lower terminal; the wiper divides the resistor in two parts whose sum is \( R \).

The circuit on Figure 4.10a is the potentiometer connection known as voltage divider. The battery voltage (\( V \)) is applied between the upper and lower ends of the resistor (\( R \)), then the current \( I_p \) in the circuit is constant \( I_p = V/R \) and the voltage at the wiper, referred to the lower (negative) terminal of the resistor, is variable. Thus, by displacing the wiper it is possible to control the output voltage of the sliding terminal between \( V \) and zero.

The potentiometer on Figure 4.10b is used as a variable resistor; the battery is connected between the wiper and one of the fixed terminals. In this circuit voltage at the sliding terminal, referred to the lower (negative) terminal of the resistor is always constant and equal to the battery voltage. Instead, the current in the circuit is variable \( I = V/[(1 - n) R] \); in practice, \( I \) is always finite because the battery cannot supply infinite current.

In many electrical devices the wiper is usually manually slipped by means of a knob or lever with the purpose of modifying the level of a voltage or current. For example, it can be used to set the volume of an audio power amplifier. Potentiometers are presented in this introduction because they can be part of electro-mechanical sensors. They can convert the motion of a mechanical device which displaces the wiper into an electrical signal. In other words, they may be a sensor, converting motion into electrical signal.

Fig. 4.10: (a) A potentiometer used as a voltage divider. (b) A potentiometer used as a variable resistor.
The concept of potentiometer as a voltage divider is used in well known devices as joysticks and touch-screen. Also they were used in tide gauges, anemometers, and instruments which need to account for mechanical displacements. At present, they are replaced by optoelectronic encoders in several applications.

**4.4 Impedance Sensors**

Devices that modify their electrical impedance due to changes in a measurand are called *impedance sensors*. They are of capacitive or inductive type, or a combination of both. The circuits used with impedances are similar to those discussed for resistive sensors (Figs. 4.1 to 4.10), but the exciting sources must have alternating waveforms. Most times they are sinusoidal voltages or currents with a frequency from kHz to MHz.

In some circuits the excitation is a voltage source whose frequency and amplitude are kept constant. The electrical output signal is produced due to changes in the impedance (capacitance or inductance). These changes are produced by the variation of the measurand that directly modifies some impedance characteristics, similar to the cases shown in Figure 4.1 for resistive sensors. Therefore, the current or the voltage is modified becoming the electrical signal output which represents the measurand.

In other circuits, the impedance is part of an oscillator whose resonant frequency is a function of the sensor’s impedance. Thus, when the measurand modifies the impedance, the oscillator’s frequency varies with the impedance changes. In this case the output signal where the measurand variation is “impressed” is the frequency. Perhaps these ideas will be better understood after reading Section (4.5).

In any case, to understand how impedance sensors work, we must study how the impedance could be changed by an external parameter such as the measurand. As shown for resistors, an impedance will also depends on its geometry (length and area) and on the material it is made of.

**4.4.1 Capacitive Sensors**

It was stated (Eq. (3.16)) (repeated in Eq. (4.6)) that the capacitance of a parallel plate capacitor is directly proportional to the facing surface of the plates ($A$) and to the permittivity of the dielectric material ($\varepsilon$) separating both plates, and inversely proportional to the distance between them ($d$),

$$C = \frac{A\varepsilon}{d}$$  \hspace{1cm} (4.6)

The capacitance can thus vary due to geometric changes or changes in the permittivity of the dielectric material. In general, for capacitive sensors, the plate area is fixed in the manufacturing process, but the facing surface between the two plates may be changed as in Figure 4.11a. Also the distance between both plates can
be modified as in Figure 4.11b. Some sensors have a fixed geometry but the dielectric material moves as in Figure 4.11c, changing the amount of dielectric between plates.

![A variable capacitor. (a) The facing surface varies. (b) The distance between plates varies. (c) The dielectric varies.](image)

In Figure 4.11, \( D \) indicates a change of distance between the plates, or in the dielectric position with respect to the plates. When the sensor is used to measure displacements, the object whose position is being measured is coupled mechanically to that part of the capacitor that produces the change in the capacitance.

In other sensors all parts are fixed but the dielectric properties of the material change due to chemical or physical changes. For example, in some humidity sensors, moisture changes the permittivity of the dielectric material.

### 4.4.2 Inductive Sensors

A variable inductance (\( L \)) is basically a winding of \( N \) turns that has a movable core of magnetic material. Equation (3.18) (repeated as Eq (4.7)) indicates how to calculate the inductance for long coils.

\[
L = \frac{\mu N^2 A}{l}
\]  

(4.7)

Theoretically, to modify the inductance, the geometry of the coil or the core magnetic permeability could be altered. In practice, the coil length and area cannot be easily modified, and then to have a variable inductance the permeability should be changed. But because the magnetic permeability is a property of the core material, once constructed it cannot be changed. Thus, the easiest way to change the inductance
is by displacing the core a given distance $D$ into or out of the coil, as shown in Figure 4.12. This device could be used to measure displacements. In general, the coil remains fixed and the core is connected to the piece which transmits the measured displacement.

**Fig. 4.12:** A movable piece attached to a magnetic core displaces it inside a coil, changing the coil inductance as $D$ varies.

Another inductive sensor is the proximity sensor (Fig. 4.13). It is composed of a coil and a magnetic core. When the coil is excited with alternating voltage an electromagnetic field is induced in the core. Any conductive piece placed close to the core will modify this magnetic field and the current through the coil will change. Then measuring the current in the coil, the presence of a conductive piece can be detected. This operating principle is used to detect the proximity of conducting materials. It is very used in industries, appliances, cars and instruments where motion has to be measured.

**Fig. 4.13:** A coil with a magnetic core is excited by an alternating voltage source thus generating a magnetic field. A conductive piece passing through the field takes energy from it “reflecting” this energy change in the coil current. Thus by measuring the coil current the conductive piece can be detected.
4.4.3 Inductive Actuators

Although they are not sensors, it seems adequate to describe this device here due to its inductive nature and because this concept is needed to explain how an oscillator works. This kind of actuator is based on the electromagnetic force, and variations of the elementary device described below are used in many applications.

As was introduced in Section (3.7.12), a coil through which a current is flowing behaves as a magnet, and it is called an electromagnet. Electromagnets are very useful as actuator devices. As explained in Section (2.1), an actuator is a device that converts an electrical signal into a mechanical motion.

The same as with transformers, practical electromagnets convey the electromagnetic field through a closed magnetic core because it greatly concentrates the effect of magnetic fields. Some cores increase the magnetic force by thousands of times with respect to the coils in air. A simplified electromagnet is depicted in Figure 4.14 where the coil and the core are shown. In practical electromagnets there is also a hollow ferromagnetic cylinder covering the coil (not shown in Figure 4.14 for the sake of clarity) which improves the efficiency of the electromagnet.

![Diagram of an electromagnet](image)

**Fig. 4.14:** An electromagnet is comprised of a coil wound on a magnetic core and a voltage source \( (V) \) which can be DC, AC or pulsed. The electromagnetic force moves the ferromagnetic piece a distance \( D \). According to the application, the ferromagnetic piece may represent a part of a mechanical actuator or several tons of steel; the theory behind both is the same.

When the electromagnet is supplied with a DC voltage it works just as a magnet, attracting ferromagnetic materials. These electromagnets are very used in equipment we use everyday, as a door lock. They permit a mechanical device to be operated from an electronic circuit. Special electromagnets are able to hold ferromagnetic pieces of several tons and they are used to move big pieces in steel industries.

When the electromagnet is supplied with a pulsed voltage, a transitory force is produced that pushes or attracts a piece; this is the kind of application we will see
later in tuning forks. If an AC voltage is applied, a periodical force is generated. This is the case of a doorbell: people pushing the button at the entrance close a circuit which, by means of AC current, makes the electromagnet vibrate a hammer over the bell.

### 4.4.4 Transformer Sensors

The most widespread transformer sensor is the Linear Variable Differential Transformer known as LVDT, which is used to measure small displacements.

Figure 4.15 shows the components of an LVDT schematically. The transformer consists of one primary (P) winding and two secondary windings (S1 and S2), P being centered between S1 and S2. The coils are wound on a hollow cylinder made of polymer. The transformer has a movable core (C) made of magnetic material which displaces axially inside the coil’s hollow. The object whose displacement is to be measured is attached to the core which slides smoothly inside the coil (no physical contact between both is required).

![LVDT diagram](image)

**Fig. 4.15:** A Linear Variable Differential Transformer (LVDT). (a) The primary coil P generates a flux Φ which is coupled equally to both secondary coils (S1 and S2) through the magnetic movable core, thus the output is null. (b) The displacement of the core to the right increases the coupling with S2, decreasing it with S1. The differential voltage (algebraic sum of both coil outputs) reflects this change. (c) The shift to the left produces the opposite effect and the differential voltage changes sign.
An alternating voltage of constant amplitude is connected to the transformer’s primary (P) and the voltage difference between the secondary coils is taken as the output signal. For practical use, this difference is generally amplified and conditioned by an electronic circuit. When the core is at the middle point of the transformer, it magnetically couples the primary magnetic flux (F) to both secondary coils in a similar way (Fig. 4.15a). Then, both secondary voltages are similar giving a zero voltage difference. As explained for transformers, the core confines and guides the magnetic field. Then when the core is shifted to the right or to the left (Figs. 4.15b and c), it guides the magnetic field towards S2 or S1, respectively. Therefore, the output has the maximum values at (b) and (c). According to the phase between the input and output signals one of the maximums is considered negative while the other positive.

LVDT are typically used to measure displacements. They are manufactured to work in ranges from ±1 mm to ±100 mm having an accuracy of 0.1 μm at the lower end of the range.

4.5 Oscillators

All timers and clocks used nowadays have some kind of crystal oscillator, and some sensors based on piezoelectric transducers are part of an oscillator circuit. Therefore, a conceptual approach to these devices follows.

4.5.1 Introduction to Oscillators

In electronics engineering an oscillator is a circuit that produces a periodic signal output, often a sine wave. In order to introduce this subject to people without an electronics background, the tuning fork oscillator example has been selected (Fig. 4.14). It is an old circuit, simple to understand.

As a first step towards the fork oscillator let us consider a mechanical pendulum clock; the pendulum has its own oscillating frequency which depends mainly on the pendulum length. The pendulum oscillation may be initiated by separating the pendulum mass from its lowest potential energy position, giving it an initial swing. Because of friction the amplitude of the oscillation will decrease, and after a while the pendulum will stop if enough energy to overcome friction is not supplied to it. This energy must be supplied by a gentle push on the pendulum, in the proper direction and time. It means that for the pendulum to continue oscillating, the energy has to be supplied in a way synchronized with the pendulum movement.

The tuning fork, likewise the pendulum, has its own oscillating frequency (which is a function of its length and mass) and, if it is desired to keep it oscillating, some external energy will also be needed to compensate frictional looses which tend
to decrease the oscillation amplitude. This energy has to be supplied also in a synchronized mode to keep the fork moving.

Figure 4.16 illustrates a two prongs tuning fork with the minimum components required to keep the fork oscillating: an excitation coil, a pick up coil and an amplifier.

![Excitation coil](image1.png)

**Fig. 4.16:** When excited by a stroke a tuning fork oscillates at its own oscillating frequency. To keep it oscillating the pick up coil takes information about the oscillation as an inductive sensor can do. This information is used to compensate the energy lost by friction through the excitation coil which works as an electromagnet.

A coil, through which a current flows, works as an electromagnet which attracts or repels ferromagnetic materials. Then if the fork is made of a ferromagnetic material a coil excited by a current could be used to excite the fork mechanically. Therefore, the excitation coil has the task of transforming electrical energy into mechanical energy. In other words it works as an actuator, supplying the tuning fork with the necessary push to keep it vibrating. The push should have enough energy to compensate the energy lost by friction and has to be applied in due course.

The pick up coil captures the vibration of the right prong as in the case of the proximity sensors (Fig. 4.13), producing an electric signal of the same frequency as that of the fork oscillation. The oscillation waveform of the prongs is used to synchronize the delivery (through the excitation coil) of the energy needed by the tuning fork to remain vibrating. Through the feedback loop, the pick up coil together with the amplifier synchronize the needed push to the left leg at the exact moment and adequate direction. Thus, the fork keeps oscillating and the electrical signal at the oscillator output may be used for timing purposes. Assuming that the fork has an oscillating frequency of 1000 Hz, a clock with a period of 1 ms can be built. The energy to keep the system working comes from an electric power supply (not shown to keep the figure legible) that feeds the amplifier.
One problem of these oscillators, used in the first electronic timing devices, is that due to temperature the length of the fork can change, thus modifying the oscillating frequency. Also, because of the size of the fork the oscillating frequencies are in the audio frequency range (low frequency).

A simplified and more general schematic of an oscillator circuit is shown in Figure 4.17. The power supply (PS) provides the electrical energy to keep the system oscillating. The triangle denotes the amplifier and the box is a mechanical and electrical filter composed of the two coils and the fork. The filter transforms electrical energy into mechanical energy and vice-versa, and works at the tuning fork oscillating frequency (which is a function of its mechanical characteristics). Thus, if a mechanical filter more stable in frequency and of a smaller size (higher frequency) were available, we would have a clock with a better resolution, more precise and accurate.

Fig. 4.17: Schematic of a tuning-fork or a piezoelectric oscillator circuit with an amplifier and a power supply.

### 4.5.2 Piezoelectric Oscillators

We will call piezoelectric oscillator any oscillator that uses the piezoelectric effect, based either on quartz crystal or on piezoelectric ceramics, because both have the same operation principle. Piezoelectric oscillators follow the same concept described for the tuning fork and can be represented also by the schematic of Figure 4.17. There is a mechanical vibration as well which is sustained by a synchronized supply of electrical energy. The vibrating element is a piezoelectric crystal as those depicted in Figure 3.30. Recall that piezoelectric crystals have two electrically conductive plates (electrodes) and they transform electrical energy into mechanical energy and vice-versa; as the mechanical and electrical filter described above does. As stated in
Section (3.8.1), the oscillating frequency of piezoelectric crystals depends on their mechanical characteristics, mainly the thickness of the crystal which defines the resonance frequency. Therefore, it could be admitted intuitively that the piezoelectric crystal could replace the tuning fork with some advantages.

Conceptually the process could be described as follows. When the crystal is exposed to a mechanical shock, a mechanical vibration is initiated at the resonant frequency of the crystal and an alternating voltage of the same frequency than the vibration will appear on its electrodes. This signal is amplified and feed backed to the electrodes of the piezoelectric crystal. Thus the oscillation is sustained because the energy delivered to the crystal compensates the losses.

In fact, piezoelectric crystals replace the tuning fork with many advantages. Piezoelectric crystals need little external energy to oscillate; their oscillating frequency is very stable (does not change with temperature); they are small and of low cost. Because they are small and the oscillating frequency depends on their size, very high frequency oscillators can be constructed. Very stable oscillators for several megahertz can be manufactured at a very low cost.

4.5.3 Piezoelectric Oscillator Applications

The piezoelectric crystal properties define the stability of the oscillator frequency. Significant research efforts have been done to build crystals with certain desired properties. Some quartz crystals are very stable and they are used as the time base for clocks in almost all time keeping devices.

Conversely, some piezoelectric crystals have been manufactured in such a way that their resonant frequencies be sensitive to some desired parameter such as temperature or force. Thus, when these crystals are employed as part of an oscillator, the oscillator output frequency will contain information on those external parameters. This is the working principle of several sensors in which the measurand modifies the resonant frequency of an oscillator.

4.6 Strain Gauges (or Force Sensors)

They are resistive sensors in which the dimensions of the resistor change under the action of a force. To understand the working principle of a strain gauge suppose that the resistor shown in Figure 3.18 is fixed at one end and a force is applied to the other, stretching it; then, the cross sectional area $A$ will decrease and the length $l$ will increase. If the deformation is small enough such that the material is within its range of elasticity, when the force is removed the material will return to its initial dimensions. The resistance of such a resistor will increase when stretched (Eq. (3.12)) and will come back to its initial value when the force is released. Unfortunately for
short resistors as that shown in Figure 3.18, the change in resistance is very low and very difficult to measure.

This property by which resistors change their values when subjected to a force is exploited to build sensors that allow a mechanical stress to be transformed into an electrical signal. A particular resistor designed to maximize the change in resistance when stretched is shown schematically in Fig. 4.18. A long wire of small diameter is bonded on an elastic material such that most of its length is located along the same direction, which will be called the sensing direction. When this device is stretched in the sensing direction, changes in the cross sectional area and length of the wire will increase its resistance.

When the ends of the wires are connected to a battery by means of two electrical cables (Fig. 4.18), and a variable force is applied in the sensing direction, changes in resistance produce a variable current. Then, a sensor that converts force signals into current signals is obtained. It is worth noting that this is just a demonstrative circuit, usually a Wheatstone bridge is used to detect the resistor changes (Fig. 4.2a).

The wires are very fragile and difficult to handle, so they are placed on an elastic backing material used to support them. This support is made of a dielectric material (usually plastic). Backing and wire form a device known as strain gauge.

When it is desired to measure the bending of a beam of a bridge or the compression of a column, a strain gauge is rigidly bonded to the specimen whose strain is being measured. The sensing direction of the wires is placed in the direction of the expected compression or extension. The backing material provides a good electrical insulation between the wires and the specimen.

**Fig. 4.18:** A strain gauge. A long wire of small diameter is disposed such that most of its length is located in the horizontal direction. It is bonded on an elastic support which, when stretched horizontally, changes the length and diameter of the wire, changing its resistance.
Today strain gauges are made in different shapes. Some have two perpendicular sensing directions. One is used to measure the desired axial deformation and the other is placed in a perpendicular direction not affected by the deformation. The second is used as part of a measuring Wheatstone bridge to compensate for changes in the resistance due to temperature, such that the Wheatstone bridge does not result unbalanced because of temperature changes. Other strain gauges have several radial sensing directions to measure in different directions simultaneously. Strain gauges are often used to study the structure of buildings and the dynamic behavior of cars, planes and machinery, and also in many instruments for environmental and hydraulic applications. Because of the ability of strain gauges to measure deformation, they are incorporated as part of some sensors such as pressure sensors.

Devices used to measure loads are called load cells. They consist of a deformable body made of a metal piece on which strain gauges are bonded. When the applied load deforms the body of the load cell, a change in the strain gauge resistance is measured by an electronic circuit which gives an output voltage. A calibration process relates the applied force to the output voltage and the transference of the load cell is established.

So far, strain gauges based in metal wires have been described, but there are also strain gauges based on semiconductor materials. Some semiconductors exhibit a piezoresistance effect which is defined as the change in electrical resistance when a stress is applied. These materials are thus used to manufacture strain gauges. The main advantage of these gauges is that they are several times more sensitive than gauges made of metals, but they are also more affected by temperature.

### 4.7 Sensors and Instruments to Measure Temperature

#### 4.7.1 Introduction

There are a large number of methods for measuring temperature. Each one has a set of applications for which it is more suitable. Some questions that should be answered when selecting a method to measure temperature are: What is the temperature range of use? Should the measurement be made in contact with or at a distance from the object whose temperature is to be measured? Which is the desired response time? Is long-term stability needed? Is high sensitivity required? What level of complexity of electronic circuits is accepted? Which is the size of the sample to be measured? What is the necessary accuracy, and which is the affordable cost? For some applications several methods could satisfy the measuring requirements, but for others only one may be adequate.

The characteristics of some commonly used sensors for measuring temperature will be described first. They are resistance temperature detectors, thermistors, thermocouples and I.C. (integrated circuit) sensors. All of them need some kind of
Sensors and Instruments to Measure Temperature

电子电路来呈现结果。有时，标准易获得的电子器件可以与这些传感器结合使用，例如数据记录器，它内含电子预处理电路来适应大多数温度传感器。这些传感器的主要特性被概述来帮助用户选择最适于其应用的传感器。

在这一节（4.7）的末尾，将介绍两种已知的红外测温方法和红外热成像。它们使用特定的光学和电子元件，并且作为完整仪器购买。它们不能像前文所述那样，使用标准电子设备。

### 4.7.2 热电阻温度传感器（RTD）

RTD是电阻型传感器，用于测量温度。它们基于金属的导电性质，即电子在金属晶格中自由移动。当温度增加时，金属晶格中的原子（离子核）振动，导致传导电子与静止晶格碰撞更加频繁，从而阻碍电子的移动。这种原子行为在宏观上表现为材料电阻率的增加。对于某些金属，电阻率与温度的函数关系在较大温度范围内（约-150到600 °C）是已知的，并且高度可重复。例如，对于铂，这种关系近似线性。

一个RTD本质上是一个由金属导线制成的电阻，其电阻随温度变化。根据式（3.12），电阻与电阻率成正比。对于RTD，电阻和面积保持不变，而电阻率随温度增加。铂和镍常用于制造RTD。铂RTD非常可重复，但价格昂贵。镍的RTD不太可重复，但价格便宜。

一个RTD的电阻值（R）作为温度（T）函数的近似表达式为

$$ R = R_0 \left(1 + \alpha_1 T + \alpha_2 T^2 \right) \quad (4.8) $$

其中，$R_0$是参考温度（通常0 °C）下的电阻，$\alpha_1$和$\alpha_2$是材料的导电性随温度变化的分数。它们是制造RTD的材料的恒定特性。例如，对于铂，$\alpha_1 = 3.84 \times 10^{-3} \ 1/°C$和$\alpha_2 = 5.83 \times 10^{-6} \ 1/°C^2$（Allocca and Stuart, 1983）。铂RTD通常被称为铂电阻温度计（PRT），并且在商用上有诸如PT100, PT500或PT1000的名称，PT后面的数字表示装置在0 °C时的电阻。铂RTD的实用范围是

$$ ( )^{210} \ \alpha_1 + \alpha_2 = $$
from -200 to 650 °C. Because of the low resistance of PT100, the resistance of the cables connecting the device to the electronic circuit could introduce errors; therefore, a four wire technique should be used (Fig. 4.4).

When it is desired to know the temperature of a given environment, RTD have to be placed in contact with such environment, and some time is required for the sensor to reach the equilibrium temperature. In general, RTD have a slow response time, for example, an industrial PT100 has a response time of 3 s in liquid and 15 s in air. Then they are not suitable to measure fast changes in temperature.

The sensitivity of a PT100 sensor is low, for a 1 °C change in temperature its resistance will vary 0.384 Ω. A large current through the sensor could produce a good voltage signal but it would also cause heating of the sensor. At the same time, heating introduces error because the RTD would report a higher temperature than that of the environment whose temperature is to be measured; this behavior is known as self heating, and errors due to it are larger when measuring in gases, due to the lower heat dissipation from the sensor.

Generally currents on the order of 1 mA are used to avoid self heating, and then signals of 384 μV /°C are obtained. Because of these low signal levels it is necessary to avoid noise entering the circuits to have and adequate signal to noise ratio. In this regard, it is necessary to avoid long cables, and to place cables away from devices that may emit electrical noise such as motors. The use of shielded cables is recommended.

In summary, RTD have large temperature ranges, high repeatability but low sensitivity and slow response time.

### 4.7.3 Thermistors

Thermistors are a kind of resistive temperature sensors made of semiconductor materials. This gives a greater sensitivity to them but their resistances vary nonlinearly with temperature.

There are two types of technologies; one of them produces devices with positive temperature coefficient (PTC) - resistance increases with temperature as in the RTD case. They are frequently used to limit current and protect electrical devices.

The other technology produces thermistors with negative temperature coefficients (NTC); these being the most used as temperature sensor for environmental applications. The transfer function for NTC thermistors has the form:

\[
R(T) = R(T_0) \exp \left[ \beta \left( \frac{1}{T} - \frac{1}{T_0} \right) \right]
\]

where \( R \) is the value of the thermistor’s electrical resistance (in Ω) at the absolute temperature \( T \) (K); \( R(T_0) \) is the resistance at a reference temperature \( T_0 \) (generally 298.15 K or 25 °C); and \( \beta \) is a constant value, characteristic of the thermistor’s manufacturing process. Manufacturers do not specify each of the thermistor
characteristics, but they provide average values of $\beta$, $R(T_0)$ and $T_0$ and their respective dispersions. Then, if users want to have the best approximation to the transfer function, they must measure $R_1$ and $R_2$ at two different known temperatures $T_1$ and $T_2$ to obtain $\beta$ for their particular sensor. Besides, generally, thermistor transferences vary with time and the sensor requires periodical calibrations to keep it accurate.

For $\beta = 3000$ K, $R(T_0) = 1000 \, \Omega$ and $T_0 = 25 \, ^\circ C$ the transfer function of an NTC thermistor is plotted in Figure 4.19. It is clearly more sensitive at low temperatures. Then, due to this nonlinearity, if sensor’s users want a better knowledge of the transference, they should measure several calibration points along the curve and fit the function by a least squares method.

![NTC Thermistor Transfer Function](image)

**Fig. 4.19:** NTC thermistor transfer function.

Many circuits have been developed with the purpose of compensating the nonlinearity of thermistor transferences, but a simple way to obtain the temperature as a function of resistance which does not need special circuits is to use some approximations. Two frequently used approximations are (Hewlett Packard, 1980; http://www.princeton.edu)

$$
\frac{1}{T} = A + B \ln R + C (\ln R)^3; \quad T = \frac{B}{\ln R - A} - C
$$

(4.10)

where $T$ is the absolute temperature (in K), $R$ the resistance of the thermistor (in $\Omega$, at temperature $T$) and $A$, $B$ and $C$ are constants to be found by selecting three $(R_i, T_i)$
data points and solving the three resultant simultaneous equations. The first of these approximations is known as the Steinhart-Hart equation. For a temperature range smaller than 100 °C the equation approximates the real value with errors less than ±0.02 °C.

The second approximation is easier to compute, but to produce similar errors to those from the Steinhart-Hart equation it should be used in a narrower temperature span (for example to measure the temperature of water in the sea, where the span is generally of a few Celsius degrees).

It is easy to find thermistors with high resistances; hence the resistance of the cables connecting them to the electronic circuit is not as important as in the RTD case.

Because thermistors are made of semiconductor materials their maximum measuring temperature is limited to about 200 °C; prolonged exposures of them close to the maximum operating limits will cause the thermistor to drift out of calibration. Manufacturing processes allow thermistors to be made with a small thermal mass; thus response times smaller than those of RTD are feasible.

Figure 4.20 shows at the center a metal encapsulated thermistor and on the right side a ceramic bed thermistor in which the sensing element is only the spherical tip, the metal wires being covered by an insulating tube (a pen is included for comparison purposes).

![Fig. 4.20: Two types of thermistors (the scale is given by the pen).](image)

## 4.7.4 Thermocouples

### 4.7.4.1 Thermoelectric Effect

It is a reversible effect by which thermal energy is converted into electrical energy or vice-versa. The first phenomenon is known as the Seebeck effect, and the second as the Peltier effect. More specifically, temperature differences are converted into voltage (or current) and backwards. Both effects have practical use in measuring temperature or changing the temperature of objects.
4.7.4.2 Thermocouple Description

Thermocouples are temperature sensors based on the thermoelectric effect. They are composed of two wires of different metals (such as constantan and copper) joined at two points (junctions). When one of the junctions is heated, a current flows through the wires due to a thermal emf (ε). Figure 4.21 shows a device which converts a temperature difference into a current. In this example the current is measured by an ammeter whose input impedance is almost zero. The current is proportional to the temperature difference between both junctions ($T_1 - T_2$).

If the two copper wires are connected to a voltmeter with a great resistance a weak current will flow, but a thermally induced electromotive force $\varepsilon$ will be measured by the voltmeter. This emf ($\varepsilon$) is proportional to the temperature difference between both junctions ($T_1 - T_2$). The proportionality constant ($k$) (Eq. 4.11) depends on the type of metals the thermocouple is made of.

$$\varepsilon = k (T_2 - T_1)$$ (4.11)

Fig. 4.21: A constantan-copper thermocouple subjected to a temperature difference between both junctions.

Usually the thermally induced emf generated by the Seebeck effect is very small. The most used metals to construct thermocouples are copper and constantan. At room temperature they produce an emf of 40 μV per degree centigrade of temperature difference between the two junctions.

Because the voltage measured is proportional to the difference in temperature, if the temperature $T_1$ is desired, it would be necessary to know $T_2$. One way to force a known temperature $T_2$ is to place the cold junction in an ice bath, then $T_2 = 0^\circ$C
becomes the reference temperature. Because the use of an ice bath is not a practical solution, an electronic compensating circuit known as electronic ice point reference is used to measure the temperature $T_1$.

Thermocouples have a wide temperature range, for example, those made with wires of platinum and platinum-rhodium can measure over 1500 °C. Thermocouples are not linear in all the measuring range, but because they are very well known there are standard tables that allow the temperature to be known by measuring the thermocouple output voltage. Also, because they have small masses they can respond fast to changes in temperature.

4.7.5 I.C. Sensor

The I.C. letters stand for integrated circuit. These are sensors based on certain properties of semiconductor junctions, and for this reason they are also known as semiconductor sensors. A semiconductor forward-biased PN junction has a linear voltage ($V_{pn}$) to temperature ($T$) relationship; for a silicon diode this relationship is given by

$$\frac{\Delta V_{pn}}{\Delta T} \approx -2.3 \text{ mV}^\circ\text{C}$$ (4.12)

This relationship is exploited to manufacture temperature sensors. Because these sensors are made of semiconductor materials the measuring temperature range is limited to -55 to 150 °C. The response time is slow due to their thermal masses. Today, these kinds of sensors have reached such a degree of development that they can be easily used, even by people with little electronic skill.

There are basically two types of I.C. sensors: with analog output and with digital output. Sensors with analog output require a minimum amount of external components to become a thermometer (National Semiconductors, 2000). In some cases only a power supply and a resistor is enough to have an analog voltage proportional to temperature. This voltage can be read on a simple voltmeter or recorded by a data logger.

Some specifications for analog output I.C. sensors are: accuracy of ± 0.25 °C at room temperature, and of ± 0.75 °C over the full range (-55 to 150 °C); they have a self-heating less than 0.1°C in still air and are suitable for remote applications.

I.C. sensors with digital output can be read directly from microprocessor standard digital buses such as F/C (Texas Instruments, 2009). Then, a user with some skill in programming could use this technology to assemble a network of temperature sensors. Some I.C. allow several devices to be networked in parallel in just one bus (few cables) No external components are required and accuracies of 0.5°C in the range from 0 to +65 °C, or of 1.0°C in the range from -40 to +125 °C, are available. Figure 4.22 shows an analog sensor on the right and a modern digital one at the center.
It should be noted that manufacturers specify the accuracy that the user could expect using the factory calibration. This is the accuracy that we can suppose from a temperature measuring instrument when the sensor has to be changed for another coming from the factory.

But because in some sensors the resolution is 0.0625°C, the accuracy could be increased several times if the sensor is previously calibrated by the user. In this case, if the sensor has to be changed, the new one has to be recalibrated again by the user in order to keep the improved accuracy. If not recalibrated, the accuracy will be that specified by the manufacturer.

Digital I.C. sensors are a good option when measurements of slow temperature changes at several points in a reduced volume are required such as in the case of a silo for grain storage or a greenhouse.

### 4.7.6 Infrared Thermometers (IRT)

Infrared thermometers, also known as infrared pyrometers, measure the surface temperature of an object at a certain distance from it. IRT are useful in applications where contact measurements are impracticable as in the case of melting metals or dangerous areas. They measure the electromagnetic radiation emitted by the object due to the object’s temperature. Some ideas about electromagnetic radiation will be introduced to facilitate the understanding of this measuring method.

A blackbody is an idealized object that absorbs all radiation arriving at it and emits a specific spectrum of energy. The intensity of the radiated energy at any particular wavelength increases as a function of temperature. The total power emitted by a black body per unit area of its surface is proportional to the fourth power of its absolute temperature (Stefan’s law).

\[ E \alpha T^4 \]  

(4.13)
where $E$ is the power per unit area ($\text{W/m}^2$) and $T$ is the absolute temperature ($\text{K}$). Therefore, by measuring the emitted energy of a blackbody it is possible to know its temperature.

Most materials do not behave as blackbodies but their temperature may be estimated by comparing their radiation to that of the blackbody. In this respect correction factors are used to relate the radiation curves of real objects to that of a blackbody. These relations give origin to the concept of emissivity which is the fraction of energy being emitted by a given surface, relative to that emitted by a blackbody surface, when both surfaces are at the same temperature.

A blackbody is a perfect emitter of heat energy, it emits all the energy that it absorbs so it is said that it has an emissivity of 1. On the contrary, a perfect thermal mirror will reflect all the energy that reaches it and is thus said to have an emissivity of 0. Most objects are neither blackbodies nor thermal mirrors; so they have emissivity values between 0 and 1. In order to obtain accurate temperature measurements from objects that are not blackbodies it is required to know the emissivity of the object being measured.

It could be difficult to know the emissivity of some surfaces whose temperature is to be measured. This hinders the exact temperature to be known by an IRT. Moreover, when there are several objects in the field of view of the IRT and each one has a different emissivity it is not correct to assign one emissivity as representative of them all.

Also, because emissivity is a property of the surface of an object, it could change with time. For example, a new polished metal surface will look shiny and will have a certain emissivity, but when it becomes corroded it will look dark and the emissivity will change. Then the same object will have different emissivities as time passes by.

Most of the emitted radiation is concentrated in a wavelength band called the infrared spectrum. Infrared energy is electromagnetic energy in a range of frequencies below the visible light; all objects radiate infrared energy not perceived by the human eye. The human eye cannot detect electromagnetic wavelengths longer than 0.7 $\mu$m (red) (http://www.x26.com/articles.html#). Some IRT are designed to be sensitive to a wavelength band around 8 to 14 $\mu$m. Very hot objects such as molten steel emit visible light and their colors are related to the temperature of the object. Visible light is a very narrow portion of the entire electromagnetic spectrum ($\approx 0.4-0.7 \mu$m).

There are different models of IRT but a generic one is pictured in Figure 4.23 to give an approximate idea on how these devices work.

Radiation from the emitting surface passes through a filter which limits the radiation wavelength range to the infrared band and focuses the energy onto a mirror. The mirror concentrates the energy on a detector which could consist of a blackened disc that heats up due to the received energy. The temperature of the disc can be measured by some temperature sensor such as a micro thermocouple or a small thermistor. As explained above, these sensors have not a very fast response; then, when it is desired to detect the temperature of fast moving objects, the detector should be a photomultiplier tube.
A photomultiplier is a device designed for the detection of photons. It exploits the secondary emission of electrons in a vacuum tube. When photons strike the tube’s cathode, electrons are emitted and attracted to a first positively charged electrode (anode). When electrons collide with the first electrode (or anode), more electrons are released and attracted to a second positively charged electrode (or anode). This process is repeated several times until there is an easily detectable current flow through the last anode.

IRT are not intended to measure temperature exactly, but only approximately. Fortunately, for most applications it is only necessary to compare the temperature of different objects, not to know their absolute values.

For example, if it is desired to know the temperature of a high voltage line contacts, all contacts would probably have a similar unknown emissivity. Then, the absolute temperature of the contacts will be measured with error because the actual emissivity is ignored. But if one of the contacts is hot due to a defective connection, the increase in its temperature relative to the others will be easily detected.

IRT have a wide temperature range of use, for example, some commercial equipments have ranges from -20 to 1000 °C for wavelengths from 8 to 14 μm, and others from 550 to 4000 °C for wavelengths from 0.85 to 1.1 μm. The first are used in ceramics, rubber, paper, food, asphalt and plastics industries; and the second in steel, glass, induction heating and forging industries. Instruments with a response time between 10 ms to 1.5 s are available.

IRT are a promising tool in environmental sciences applications where it is required to measure a considerable area from a certain distance. It has been found that low-cost consumer-quality IRT are useful for non-contact measurement of plant canopy temperature (Mahana & Yeaterb, 2008). They can also be used for studies
on land-atmosphere interaction, surface energy balance, human animal body temperature measurements, etc.

Because the energy radiated from an object whose temperature is being measured is transmitted through the air and received by the thermometer, the results may be affected by smoke and dust. Thermometers receive radiation caused by the ambient temperature in the room together with the radiated energy from the object being measured; to avoid measuring undesired background temperatures, the infrared energy collected by the thermometer should be originated by the particular area whose temperature is measured.

### 4.7.7 Infrared Thermography

Infrared thermometers measure the *average temperature of an area*. Instead infrared thermography uses infrared cameras which detect energy emitted from individual points of an object, convert them into temperature data, and display an image of temperature distribution. Thermography converts thus infrared radiation into a visual image of the temperature distribution similar to that shown in Figure 4.24.

![Fig. 4.24](image)

**Fig. 4.24:** (a) Photography of a drying leave. (b) The same leave as seen in a thermographic image.

Different temperatures on a surface are represented by different colors and temperature comparisons over a large area are possible. This permits hot spots to be found quickly. A temperature image consists of a matrix of pixels and the number of pixels depends on the number of detectors. Some thermal cameras have for example $320 \times 240$ pixels which are colored up pixel by pixel by the onboard software according to the temperature measured by each detector.
This technology lets thermal anomalies of rotating equipments, such as big engines to be detected by periodical inspection under the same running conditions. It also permits safe inspection of live electrical equipment difficult to access. Food temperature can be measured in a sanitary fashion.

In environmental sciences, thermography allows semiautomatic analysis of large areas of canopy with effective replication of measurements (Jones et al., 2002). It has also been demonstrated the system’s ability to survey populations of several wildlife species. It permits different species in the same habitat to be detected, and it is expected that with the aid of computer-assisted analysis, infrared thermography may become a useful wildlife population survey tool (Garner et al., 1995).

### 4.7.8 Comparison of Temperature Sensors and Instruments

In order to select the most adequate temperature sensor for a given application, their general characteristics are shown in Table 4.2.

**Table 4.2: Characteristics of temperature sensors**

<table>
<thead>
<tr>
<th>Temperature sensor</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RTD</strong></td>
<td>Most repeatable</td>
<td>Small sensitivity</td>
</tr>
<tr>
<td></td>
<td>Accurate</td>
<td>Low resistance</td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>Slow response</td>
</tr>
<tr>
<td></td>
<td>Wide temperature range</td>
<td>Self heating</td>
</tr>
<tr>
<td><strong>Thermistor</strong></td>
<td>High sensitivity</td>
<td>Non linear</td>
</tr>
<tr>
<td></td>
<td>Fast response</td>
<td>Limited temperature range</td>
</tr>
<tr>
<td></td>
<td>High resistance</td>
<td>Fragile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self heating</td>
</tr>
<tr>
<td><strong>Thermocouple</strong></td>
<td>Wide temperature range</td>
<td>Non linear</td>
</tr>
<tr>
<td></td>
<td>Robust</td>
<td>Small sensitivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature reference needed</td>
</tr>
<tr>
<td><strong>I.C. sensor</strong></td>
<td>Linear</td>
<td>Limited temperature range</td>
</tr>
<tr>
<td></td>
<td>High sensitivity</td>
<td>Slow response</td>
</tr>
<tr>
<td></td>
<td>Inexpensive</td>
<td>Poor factory calibration</td>
</tr>
<tr>
<td></td>
<td>Analog and digital output</td>
<td></td>
</tr>
<tr>
<td><strong>Infra Red Thermometer</strong></td>
<td>Wide temperature range</td>
<td>Not accurate</td>
</tr>
<tr>
<td></td>
<td>Measure at distance</td>
<td>Periodical calibration required</td>
</tr>
<tr>
<td><strong>Infra Red Thermography</strong></td>
<td>Same as IR thermometer</td>
<td>Same as IR thermometer</td>
</tr>
<tr>
<td></td>
<td>+ Image of temperature</td>
<td>Cost</td>
</tr>
<tr>
<td></td>
<td>Distribution</td>
<td></td>
</tr>
</tbody>
</table>
4.8 Pressure Sensors

4.8.1 Introduction

Pressure sensors give an electrical signal output proportional to the applied pressure. Figure 4.25 shows a device to convert an applied pressure into a voltage signal. This is a schematic of an idealized device which meets certain characteristics common to various technologies used for pressure sensors. This schematic allows a simple explanation on how these sensors work.

The device comprises a circular membrane placed on the top of a rigid cylinder. The applied pressure bends the membrane made of an elastic material such as a thin metal. While the material is in the elastic range, the membrane will return to its initial conditions when pressure is released. The material, shape and size of the membrane are selected to have a deformation proportional to the applied pressure.

One way to account for membrane deformation is by means of a strain gauge (Section 4.6). Then, bonding a strain gauge to the elastic membrane, connecting it to a Wheatstone bridge (Fig. 4.2) and the bridge to an amplifier, a voltage output proportional to the applied pressure is obtained.

In summary, pressure deforms the membrane and the deformation is measured by a strain gauge which converts deformation into an electrical signal. Therefore, by determining the relationship between the applied pressure and the measured voltage the transfer curve of the sensor can be established.

It should be pointed out that the membrane is free to deform and that the only restoring force is due to the elastic force of the material. No force is applied to the back of the membrane. A reference port is depicted in Figure 4.25 which vents the reference chamber located inside the cylinder to the atmospheric pressure, so that air is not compressed by the displacement of the membrane.

![Fig. 4.25: Schematic of a device for converting an applied pressure into a voltage signal.](image)
4.8.2 Types of Sensors

4.8.2.1 Differential Sensors

One way to know the difference between two pressures is to measure both with two independent sensors such as that previously described. Each one will give an output voltage proportional to the pressure to which it is subjected. Each membrane is referred to the atmospheric pressure through the reference port. The voltage difference that results from subtracting both outputs is proportional to the pressure difference.

A more accurate method for measuring pressure differences than that explained above follows. The membrane of the sensor of Figure 4.25 can receive pressure from the top and the bottom. Then, if the upper side of the membrane is exposed to one of the pressures to be measured and the reference port to the other, the membrane will deform according to the pressure difference. Thus, the strain gauge with the associate circuit will give an output voltage directly proportional to the differential pressure.

Manufacturers of pressure sensors utilize a particular classification that users should know to correctly select sensors for diverse applications. Sensors that have two pressure ports accessible to users (as in this case) are called “differential sensors”. The pressure port related to the upper side of the membrane is called the measuring port and that related to the back of the membrane, the reference port.

4.8.2.2 Vented Gauge Sensor

When measuring fluid levels with a pressure sensor, the pressure that receives the measuring port is that of the fluid column plus the atmospheric pressure. Therefore, if the reference port is left open to air, it will sense the atmospheric pressure and the membrane will measure the difference, compensating for atmospheric pressure changes. This is a particular case of differential pressure sensors where the reference port is vented to the local atmospheric pressure. This kind of sensors is known as “vented gauge sensors.” This is the case depicted in Figure 4.25.

In most industrial processes it is quite simple to measure fluid levels with a vented gauge sensor because, in general, the pressure exerted by the fluid column is connected to the measuring port by means of a tube and it is not difficult to expose the reference port to atmospheric pressure, but it could be particularly complex in other circumstances and precautions should be taken.

For example, some manufacturers offer submersible vented pressure sensors to measure water level in lakes, rivers and harbors. In order to expose the reference port to the atmospheric pressure, they utilize a vent tube into the same cable used for the electric power supply and output signals. This vent tube would allow atmospheric pressure variations to be compensated, but in practice any obstruction restricting air movement along the vent tube could prevent the atmospheric pressure from reaching the sensing membrane and would introduce errors in the measurements. Sometimes humidity condensation is the cause of the vent tube obstruction.
Because humidity in the tube can cause inaccurate readings, a desiccant must be installed and periodically replaced to avoid condensation. This routine is not always possible and sometimes unpractical, as it could be the case of unattended instruments. According to the particular application, it should be careful evaluated when it is convenient to use this kind of sensors with compensating vented tube.

**4.8.2.3 Sealed Gauge Sensors**
These sensors have the reference port hermetically sealed and the reference pressure chamber is at a certain pressure within the atmospheric pressure range. This pressure is defined by the manufacturer at the moment of sealing the reference port. Sealed gauge sensors are useful to measure dynamic pressures and allow positive and negative pressures to be measured with respect to the sealed atmospheric pressure.

When sealed gauge sensors are used to measure water levels, the varying atmospheric pressure is superimposed to the water column pressure and may cause measuring fluctuation on the order of ±0.1 m. Therefore, to evaluate the exact level it is needed to measure and subtract the atmospheric pressure.

**4.8.2.4 Absolute Pressure Sensor**
If a vacuum is created into the reference pressure chamber and then the port is sealed, an absolute pressure sensor is obtained. In this case the external pressure on the measuring port is relative to vacuum.

**4.8.3 Other Pressure Sensor Technologies**

For most pressure sensors, as those described above, the measuring pressure produces a mechanical stress on the sensor which changes some electrical parameter in it. In the previous case the changing parameter was the resistance of a strain gage bonded to a membrane. In other sensors the mechanical stress may change the properties of a semiconductor or the frequency of a resonant circuit.

**4.8.3.1 Solid State Pressure Sensors**
In solid state pressure sensors the elastic membrane and the variable resistances are integrated in a small wafer of silicon where the sensing element and some additional circuit components are mounted on a ceramic substrate. They are known as IC (Integrated Circuit) pressure transducers and some manufacturers package them in a dual-in-line configuration (two parallel rows of pins, similar to that of IC), then sensors may be mounted on printed circuit boards. Often, these sensors require only power supply to provide a voltage output proportional to pressure.
Silicon micromachining techniques are used to implant piezoresistive strain gages into a Wheatstone bridge configuration on board the chip. Because massive production techniques are used, these sensors are of low cost. Some sensors include temperature compensation but, in general, IC pressure sensors are more sensitive to temperature changes than those from other technologies. Frequently they are used where high accuracy is not required and low cost is considered.

4.8.3.2 Quartz Crystal Pressure Sensors

When high accuracy measurements are needed, special quartz crystal resonator sensors are used. As explained above (Section (4.5.2)), crystal oscillators may have frequency outputs related to a measurand. It is only required that the measurand change the resonant frequency of the crystal. In the pressure sensor case, by means of certain special designs, crystal frequency is a function of the stress exerted on the crystal. The resonant frequency output is sustained and detected with electronic circuits similar to those used in precision clocks, as was previously described (Section (4.5.3)).

Different types of quartz crystal stress-sensitive resonators have been developed (http://www.paroscientific.com/qtechnology.htm). One of them consists of two similar beams with two mounting pads, on which stress is applied (Fig. 4.26a). The beams are electrically excited by a pair of electrodes. Frequency is a function of the applied stress, increasing with tension such as it happens with a guitar string.

Until now we have a stress sensor, but we want to measure pressure; then, by means of some mechanical device, pressure has to be transformed into stress and the stress applied to the crystal (Fig. 4.26b). In this schematic the quartz crystal sensor is fastened at the bottom to a fixed plate and aneroid bellows transform the input pressure into a deformation, such that a force is generated and applied to the other end of the sensor.

This kind of sensors can achieve 0.01 % accuracy, 0.0001 % resolution, and have high reliability and stability; their output frequency makes them less susceptible to interference. Also, they are easy to interface to counters, microprocessors and computers. This kind of technology is more expensive than those previously described.

One of the manufacturers of this kind of sensors claims that a sensor installed at a depth of 6,000 meters was able to detect an earthquake-generated wave (tsunami) (Yilmaz et al., 2004). The real signals were resolved to 1 mm of water (1 part in 6 million) and clearly show the characteristics of the tsunami which was only several centimeters at the deployed depth of thousands of meters. This pressure transducers have accuracy better than 100 parts per million of full scale pressure and maintain this accuracy for a long time.

Long-term stability tests indicate that the median drift rate of three units, tested during a fifteen-year test period, is 7 parts per million per year, which is an excellent figure.
**Fig. 4.26:** (a) Two similar beams with two mounting pads on which stress is applied; beams are electrically excited by a pair of electrodes. (b) Aneroid bellows transform pressure into a deformation, such that a force is applied to the sensor.

### 4.8.4 Application of Pressure Sensors to Measure Water Level

When applying pressure sensors to water level measurements it is convenient to evaluate some constraints. A real case example will be introduced with the aim of weighing up different features of the problem.

It was needed to simultaneously measure water level at several points of the drainage channels of an urban area to know how they are filled during rain events. The channels run below the streets of a city and they are about 3 m high and 7 m wide.

Initially, submersible pressure sensors of the type *vented gauge sensor* were considered for this application due to their ability to compensate for atmospheric pressure changes, but some problems arose that influenced to change the first thought.

Since the vent tube of the vented gauge sensor is included in the same cable used for the electric wires, it is convenient to buy cables with the suitable length required for the installation such that not electrical and pressure connections have to be done in the field.

Because access to the channels was difficult and required logistic support from the police force and the fire-brigade, at the time the sensors had to be bought there was no information about how to run the cables from the drainage channels to the surface. Therefore, it was required to have the flexibility to define the length of the cables during the installation of the instrument in the field, which is easier done with cables that only have electrical wires.

The environmental humidity at the instrument installation places was quite high and condensation in the cable vent tube could happen. The use of silica gel at the end
of the tube requires a periodical service that was decided to be avoided because there was no possibility of a frequent visit to the instruments.

The above reasons and the lack of experience using vented tubes, added to the client’s demands to avoid collecting erroneous data, biased the decision towards the use of sealed gauge sensors. They entail the use of extra instruments to measure atmospheric pressure but they looked less prone to failures or errors. A fact that contributed to this choice was that all the installations were in a reduced area where the atmospheric pressure could be considered the same. Then only one atmospheric pressure recorder was needed to compensate all sealed sensors.

It has to be underlined that it is not intended to say that this is the best possible solution to the problem presented in this example; other valuables solutions could be chosen as well. Sometimes, the solutions are selected according to the user’s previous experiences and circumstantial constraints. Therefore, sensors with sealed pressure reference ports were selected for this application to avoid humidity access, and a simultaneous atmospheric pressure sensor used to compensate atmospheric variations.

Figure 4.27 shows the installation of a pressure sensor in a drainage channel as described before. Sensors are not installed on the floor of the channel because water carries sediment which might obstruct the sensor pressure port. Also, the sensor was placed inside a steel tube mounted on the wall to protect it from sharp and heavy objects transported by the flow. Sensors were made of stainless steel, then, to avoid galvanic corrosion they were isolated from the tube and plastic bolts were used to fix them.

Fig. 4.27: Installation of a pressure sensor in a drainage channel.
Figure 4.28 shows the water level measured in the aforementioned drainage; when the drainage is dry, as in Figure 4.27, the atmospheric pressure is the only pressure recorded. When it rains in the area, the channel begins to fill, the water surface arrives to the sensor level and it begins to record the atmospheric pressure plus the column of water over it.

Because sensors do not measure from the floor of the channel, to know the real water column the atmospheric pressure column has to be subtracted and the installation height added. Obviously, changes in the water level below the sensor’s level were not measured, but they are not data of interest in this study.

As it happen with most sensors, pressure sensors can suffer some drift on their transference as time passes. Therefore, it is recommended to perform some simple tests before spending time and money in field work. Some tests on five pressure sensors are presented on Section (11.2), with the purpose of underlining the idea that sensors are neither perfect nor invariant, and they need to be controlled periodically.

Fig. 4.28: Water level measured in the drainage channel of Figure 4.27.

4.9 Humidity Sensors

This is another example of a resistive sensor. In this case the conductivity of the material of which the resistor is made varies with the measurand. These sensors consist of two electrodes encapsulated in a block of a porous material that in the early designs was gypsum.

Contrary to what happens with strain gauges, for humidity sensors the geometry of the resistor is fixed because the size and separation of electrodes and the size of the porous block remain constant. What changes is the conductivity of the material. Initially, the block is dry and the electrical conductivity of the sensor is low because
gypsum pores are full of dry air, which is a poor electrical conductor. When the sensor is buried in wet soil, humidity will be absorbed by the porous material and air replaced by water (more conductive than dry air), thus the total conductivity of the block will increase.

![Humidity Sensor Diagram](image)

**Fig. 4.29**: Humidity sensor. Two electrodes are embedded in a porous material. When water replaces air in the porous material the electrical conductivity of the material changes. A voltage source, a resistor and an OA give an output voltage proportional to conductivity.

The transference of electrical conductivity to humidity may be obtained by a calibration process. A circuit as shown in Figure 4.29 (with two resistors in series as in Figs. 3.19a and 4.7) may be used to measure the resistance between electrodes; then, the conductivity is calculated and the humidity inferred. Most times an AC voltage is used to decrease electrode corrosion and to improve signal to noise ratio.

This kind of sensor has some drawbacks: changes in temperature and in soil water salinity modify the sensor transference; the time response of the sensor is long because it takes several hours for the block to reach soil humidity. The size of the sensor is not suited for using it in potted plants and the life of the sensors was short in early designs. At present, electrodes are made of materials which suffer low corrosion such as stainless steel, and porous materials such as fiberglass are used to give a longer life to sensors.

The same schematic of Figure 4.29 can be used to explain capacitive humidity sensors. In this case, the electrodes are the plates of a capacitor and a dielectric material is placed between the plates. The permittivity of the dielectric material (Eq. (4.6)) varies with the humidity.
Capacitance sensors are small compared with the above conductive sensors, they are flat and sizes are below 5 mm by 5 mm. The capacitance of the sensor with relative humidity (RH) of 30% is about 150 pF and a sensitivity of 0.25 pF / % RH is available. The sensor has to be excited with an AC voltage with frequencies in the range 1 to 100 kHz (http://www.ist-usadivision.com/sensors/humidity/). In general, humidity sensors measure with a relative good accuracy up to a point slightly away from both ends of the relative humidity range where the accuracy reduces.

Other capacitive air or gas humidity sensor allows measurements in the range 0 – 100% RH with errors ± 2% in the central part of the range; errors increases at the low and high 10% extremes of the range. They are also flat and small size (3 mm by 3 mm) (http://www.sensirion.com/en/home/)

Because capacitive sensors are small, the response time is fast, they can reach 63% of a step humidity change in less than 10 s.

Sensors to measure grain humidity (wheat, barley, corn, rice, soybean, etc.) are also capacitive. In this case the capacitor is cylindrical of a considerable volume (about half a liter). The grain is the dielectric material. They have a higher accuracy (0.5%) but they measure only in a reduced range (5 to 40% RH) which is the range of interest for cereals.

### 4.10 Conductivity Sensors for Fluids

Electrical conductivity of liquids allows the amount of dissolved salts to be known, a parameter of interest in oceanography, hydrology, industries, etc. Instruments to measure the electrical conductivity of liquids are composed of a sensor and an electronic circuit. The first is introduced into the liquid, whereas the second excites the sensor, process the signal and presents the results to the users. Results may be shown in a display or recorded in a non volatile memory for further analysis.

Because the conductivity of some liquids is very sensitive to temperature (e.g. for water the sensitivity coefficient is about 2 % per °C), conductivity measurements are ordinarily referred to a standard temperature of 25 °C. In order to refer the conductivity measured at the actual temperature to the standard temperature, the liquid temperature has to be measured close to the conductivity sensor. For a number of solutions, temperature coefficients are not constant and some kind of correction table should be stored in the memory of the conductivity meter to provide a means for automatic corrections.

There are two prevailing technologies for measuring conductivity; they are based on two quite different kinds of sensors: **conductivity cell and inductive probe**. Each one has their advantages and limitations.
4.10.1 Conductivity Cell

This kind of sensor uses two electrodes which are in direct contact with the liquid. Electrodes are made of different materials according to the liquid in which they will be used (stainless steel, platinum, graphite, platinum coated with platinum black, titanium, gold-plated nickel, etc). The simplest conductivity cell consists of two parallel plates (electrodes) separated by a fixed distance. This cell is shown in Figure 4.30; this figure is helpful to introduce the concept of “cell constant” and to show how conductivity units are derived.

In order to measure the conductivity of the liquid, the cell electrodes are excited with an alternating voltage, and the current through the liquid is measured (Fig. 4.30). From Eq. (3.13), if $V$ and $I$ are known, the conductance $G$ may be calculated. This conductance corresponds approximately to the conductance of the volume of liquid contained between both electrodes.

The conductance is a parameter that depends on the geometry of the electrodes and the conductivity of the liquid ($\sigma$) (Eq. (3.12b)) (shown again in Eq. (4.14)).

$$G = \frac{A\sigma}{l} \tag{4.14}$$

![Diagram of Conductivity Cell](image)

**Fig. 4.30:** Conductivity cell. The size and separation of the electrodes are fixed, but the conductivity is variable.
Let us assume that both electrode plates have an area $A = 1 \text{ cm}^2$ and are separated by $l = 1 \text{ cm}$; keeping in mind that the unit of $G$ is the siemens (S), the conductivity of the liquid may be calculated from Eq. (4.15), resulting the unit frequently used for conductivity, S/cm. The unit adopted by the International System of Units is S/m, but $\mu\text{S/cm} = 10^{-6} \text{S/cm} = 10^{-4} \text{S/m}$ is also used for many purposes.

$$
\sigma = \frac{Gl}{A} = G K \frac{S}{\text{cm}}
$$

(4.15)

The cell constant $K = l/A$ (cm$^{-1}$) represents the dimensions and shape of the electrodes. It has been adopted in conductivity literature that for a cell having two electrodes with the above dimensions ($A = 1 \text{ cm}^2$ and $l = 1 \text{ cm}$) $K = 1 \text{ cm}^{-1}$.

Conductivity meters should cover a broad conductivity range; pure water has conductivity below 0.5 $\mu\text{S/cm}$, drinking water, about 50 $\mu\text{S/cm}$, and sea water about 50 mS/cm. Therefore, to keep measuring errors due to electronic circuits as low as possible it is necessary to maintain $V$ and $I$ within a certain range (not very small values so as to keep them greater than the background noise). To achieve this purpose it is required to change the cell constant (cells of different dimensions) according to the expected range of the conductivity of the liquid inside the cell. Therefore, it seems reasonable to have different cells for each conductivity meter. For example, for low conductivity liquids, the current would result small, then the cell used to measure them should have a cell constant also low, i.e. large $A$ and small $l$. This assures to have reasonably large current which can be easily measured with small error. Table 4.3 presents three cell constants and their recommended measuring range.

Some meters have a fixed AC frequency to excite the cell, but others may allow the user to adjust the frequency to optimize signal to noise ratio avoiding noisy frequency ranges; in the last case frequency range is usually between 20 Hz and 20 kHz.

<table>
<thead>
<tr>
<th>Cell constant</th>
<th>Recommended conductivity range ($\mu\text{S/cm}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.5 to 400</td>
</tr>
<tr>
<td>1</td>
<td>10 to 2000</td>
</tr>
<tr>
<td>10</td>
<td>1000 to 200,000</td>
</tr>
</tbody>
</table>

Sometimes, with the purpose of expressing the ability of liquids to conduct electricity, the resistivity of the solution is used instead of its conductivity. The unit of resistivity is $\Omega \text{m}$. As stated in Section (3.7.5), resistivity is the reciprocal of conductivity, so values of $\mu\text{S/cm}$ correspond to values of M$\Omega$ cm.

Because electrodes are in direct contact with the liquid, measurements may result affected by different kinds of phenomena that could alter electrodes, such as fouling,
electrochemical plating and polarization. In these cases electrodes must be cleaned to renew the active surface of the cell. Abrasives or sharp objects should not be used to clean electrodes because they could be permanently damaged, but water with liquid detergent could help. A piece of cotton can be used with great caution to avoid modifying the gap between electrodes because it would change the cell constant. For most cells the cotton may be soaked in acetone or a chlorine solution. Before using conductivity meters, it is recommended that they be calibrated to a standard solution with conductivity close to the solution to be measured.

Conductivity cells are used in laboratory conditions, and frequent electrode cleaning is thus feasible in this environment, but could result impractical for autonomous instruments deployed in the field. Therefore, for field applications other type of sensors may be more adequate.

### 4.10.2 Inductive Probe

The working principle of this type of sensor is based on concepts discussed in Section (3.7.17). The sensor is composed of two toroids, each one with its own coil wound around the core (Fig. 4.31). Both toroids are submerged in the liquid whose conductivity is being measured, thus the liquid fills the central hole of the toroids. One of them, which will be called the primary, is excited with an AC voltage so that the primary current generates a variable flux $\Phi$ that is confined to the primary core. This flux produces an AC electrical potential around the core which creates an electrical current in the liquid. The conductive liquid acts like a transformer whose secondary has only one turn.

![Inductive Probe Diagram](image)

**Fig. 4.31:** Inductive probe. The primary toroid generates an electric current in the fluid proportional to fluid conductivity and the secondary measure a voltage proportional to the fluid current.
The current in the liquid acts as the primary of a second transformer and generates a flux $\Phi$ in the second toroid. This flux, which is confined to the second core, induces a voltage in the secondary coil. This voltage is the input of a measuring circuit. The higher the conductivity of the liquid the higher is the resulting current in the liquid and the secondary voltage. As usual, a transference curve that relates output voltage to conductivity is obtained by a calibration process.

Actually, when manufactured, the set of toroids together with their coils are encapsulated in materials (e.g., synthetic resin) not affected by the liquid properties (i.e. corrosivity), then they are not in direct contact with liquids, and no metal/solution contact exists. Inductive probes require low maintenance and suffer minimal effects due to fouling, thus resulting more adequate for long time deployment.

### 4.11 Accelerometers

Sensors that measure acceleration are called accelerometers. The most familiar way to make an accelerometer is with a mass and a spring mounted on a frame. The frame is attached to the body whose acceleration is measured. The displacement of the mass relative to its supporting frame is used to measure acceleration. In this sensor it is supposed that this relative mass displacement equals the frame displacement induced by the accelerated frame, as measured from outside the accelerometer. This assumption is true for a certain range of frequencies of the accelerating forces.

A net force $F$ acting on a mass $m$, subject to an acceleration $a$, deforms the spring whose displacement $x$ is proportional to the force through the spring constant $k$ (Hooke’s law) (Weinberg, 1999),

$$ F = m a \quad \text{and} \quad F = k x; \quad x = \frac{m}{k} a \quad (4.16) $$

Because $m$ and $k$ are constant, measuring the displacement $x$ allows the acceleration to be known.

There are different electrical ways to measure the relative motion of the mass and the frame, the simplest to understand is the capacitive method explained below.

Figure 4.32 shows the seismic mass, the springs and the frame. A central electrode attached to the mass, together with two external electrodes attached to the frame, forms two variable capacitors of similar initial capacitances. The position of the mass, relative to the frame is converted into an electrical signal by means of these capacitors. When the frame is accelerated upwards the lower spring is compressed, the central electrode moves downwards and the capacitances change. Let us remember that this change in capacitance produces a change in the impedance which can be measured by an electronic circuit.

Because of the upward acceleration, the distance between the capacitor’s plates are modified (Fig. 4.32). The central electrode moves downwards, thus decreasing
the distance between the plates of the lower capacitor and increasing its capacitance (Eq. (3.16)). The opposite happens with the upper capacitor.

A simplified electric circuit is shown on the right side of the Figure 4.33. An alternating voltage $V$ excites a bridge with two capacitors ($C$) and two resistors ($R$). The voltage drop between the mid points of both the resistive and the capacitive branches enters the amplifier ($A$). The bridge is initially balanced and the signal entering the amplifier is thus zero. When the seismic mass is displaced, the capacitances change due to the change in the distance between the plates. The bridge becomes unbalanced and a voltage difference between the mid points of both branches appears; this difference is amplified. The dashed rectangle relates the mechanical and electrical representation of the capacitors.

**Fig. 4.32**: Accelerometer: the mechanical set is composed of a mass, springs and a frame. The electrical sensor that measures the relative motion between the frame and the mass is a capacitor whose plates are attached to the frame and the mass.

Summarizing: acceleration produces a relative motion of the mass and the capacitor plates which is directly related to it. The change in the plate distances produces an impedance variation which is determined by a bridge, providing a voltage output directly related to the acceleration.

During the design of an accelerometer the two parameters under control are the spring stiffness ($k$) and the mass ($m$). The resonant frequency of accelerometers and the useful frequency bandwidth are established by these two parameters.
The type of accelerometer previously described belongs to the kind used for vehicle guidance, navigation and speed measurement and attitude determinations. Accelerometers for other applications require different mass and spring designs. An accelerometer similar to one used in wave measuring buoys is schematically depicted in Figure 4.34. It has a flexible beam supported at only one end (or cantilever) which accomplishes both the roles of mass and spring. The distance between the cantilever support and the fixed electrodes is constant, but when the whole system moves, the cantilever electrode shifts with respect to the fixed electrodes. As before, the capacitances between the moving and fixed electrodes change.
The diameter of the cantilever is about 0.1 mm and the gaps between the cantilever and the fixed electrodes are about 2 mm. The device is housed in a box full of liquid. The liquid plays two roles: one mechanical, damping the beam movement and another electrical, as the dielectric material of a capacitor.

Electrical wires are connected to each electrode and a circuit similar to that shown in Figure 4.33 could be employed to have a voltage proportional to acceleration. This kind of accelerometer produces a transference curve similar to that shown in Figures 2.12a and b (Section (2.4.3)).

If a conducting liquid were used, the two variable capacitors would be transformed into a potentiometer as in Figure 4.10a (or an impedance with resistive and capacitive components); the fixed electrodes being the ends of the potentiometer and the cantilever the center. The circuit of Figure 4.33 would change to a bridge with four resistors (or impedances), where the two variable capacitors are replaced by the potentiometer.

4.11.1 MEMS Accelerometers

In the last two decades Micro-Electro-Mechanical Systems (MEMS) have evolved fast resulting in many micro-machined devices. MEMS have strongly influenced the development of small sensors. One of these sensors is the surface-micro-machined monolithic accelerometer. This device includes both the signal conditioning circuitry and the sensor, manufactured together on a single monolithic chip. Conceptually they do not differ from the previously described capacitive accelerometers. The great change can be found in its small size and low cost of production.

These devices also comprise a seismic mass, springs and variable capacitors. Figure 4.35 represents the mechanical part of a MEM accelerometer where the seismic mass and central electrodes have the appearance of a fish backbone. In the figure, the central electrodes are called movable fingers and the external electrodes are called fixed fingers. For one of the commercially available MEM accelerometers (Weinberg, 1999) the overlapped finger’s length is 125 μm, the finger’s thick is 2 μm and the gaps between the central electrode and any of the external electrodes is 1.3 μm.

From an electrical point of view, electrodes play the same role than in Figure 4.33, but because each capacitor is very small, many capacitors are required to reach some minimum capacitance required to have an adequate sensitivity, then more than 60 small capacitors are summed up (Samuels, 1996).

When the seismic mass with the central electrodes are perfectly centered, both sides of the differential capacitor have similar capacitance, and the output of the sensor is adopted as zero output. However, if the seismic mass is displaced because the device is accelerating, the variable capacitor becomes unbalanced and an output signal related to the acceleration is obtained.
The electronic circuit used to get a voltage proportional to acceleration is more complex than that shown for the previous accelerometers, but because it is integrated in the chip, users do not have to care about it. Nowadays, three axes accelerometers are commercially available at low cost.

### 4.11.2 Piezoelectric Accelerometers

Some accelerometers use the piezoelectric effect, explained in Section (3.8), to generate a voltage proportional to the acceleration. Figure 4.36 depicts this sensor which comprises a pre-loaded spring, a seismic mass, a piezoelectric crystal, a base and a frame. The crystal structure is initially stressed by the spring force and acceleration forces make the mass move, which produces changes on the stress applied to the crystal. A stress change on the crystal generates a voltage on the crystal electrodes as explained in Section (3.8). Then a voltage output on the crystal electrodes, proportional to acceleration, is obtained within some acceleration's frequency range.
4.11.3 Accelerometer Applications

Accelerometers are mounted on objects, animals or vehicles to know their dynamic behavior. They produce a signal output $a(t)$ proportional to the acceleration that they experience. The transfer function of an accelerometer relates the output voltage to the input acceleration; its sensitivity is expressed in $V/g$ or $V s^2/m$; where $g$ (the acceleration due to gravity) $\approx 9.8 \text{ m/s}^2$. The velocity $v$ and the distance $x$ can be calculated from the acceleration as follows:

$$a(t) = \frac{dv}{dt}; \quad v = \int a(t) \, dt; \quad v(t) = \frac{dx}{dt}; \quad x = \int v(t) \, dt \quad (4.17)$$

When a three orthogonal axes accelerometer is used, it is possible to describe the motion of a moving body in space. Accelerometers are widely used in cars to switch the air bags, in airplanes and ships to measure their dynamic response along three orthogonal axes. In ocean sciences they are used as part of the navigation systems of remote operated vehicles. Accelerometers can be used for determining the position of moving objects floating on the sea surface and then, indirectly, the surface motion can be estimated.

In particular, due to their small size, MEM accelerometers have a wide range of applications to know motion, for example in robots or to study the dynamic behavior.
of racing cars. Nowadays, there is an increasing use of these sensors to understand the dynamic behavior of animals and human beings, for example the arms of a swimmer or the motion of a runner.

### 4.12 Geophones

Geophones are devices used to measure small vibrations passing through soils. The device depicted in Figure 4.37 represents a geophone. It comprises a seismic mass \((m)\) suspended by springs from a frame in a way similar to that of an accelerometer. Here the seismic mass is a coil of wire that moves in a magnetic field. The magnet is fixed to the frame (Oome, 2008), which is in contact with the ground or the object whose motion is to be measured. When the geophone’s frame is vertically moved, the springs are stretched and compressed and there is a relative motion between the frame and the mass. This motion is detected by the coil, which generates an electrical signal output. If there is no motion of the coil relative to the magnet, there is no electrical signal.

These geophones work in the same way as inductive microphones, where a magnet is surrounded by a coil (Fig. 4.37). When a vertical velocity is imposed on the geophone the mass moves and a voltage is generated, according to Faraday’s law (Eq. (3.23)).

Faraday’s law relates the induced emf \((\varepsilon)\) on the coil with the rate of change of flux through the coil. In the device depicted in Figure 4.37 the spatial distribution of the flux varies, therefore, when the coil moves the flux in the coil changes, thus giving a voltage at the coil terminals. Equation (4.18) shows that the output voltage \(V\) is proportional to the velocity of the frame, \(S\) being the sensitivity (in V s/m).

\[
V = S \frac{dx}{dt} \quad (4.18)
\]

This sensor in which the relative velocity between the mass and the frame is converted into a voltage by means of a coil and a magnet is called an electromagnetic geophone. In this sensor, it is supposed that the mass velocity equals the frame velocity, which is true for a certain range of frequencies.

A generic transference is plotted in Figure 4.38 where the high pass characteristic of geophones is observed. In this figure the amplitude of the signal is expressed in dB relative to the maximum.

Geophones are insensitive to low frequencies and some recent research indicates that the use of MEM accelerometers may have advantages over geophones to measure in the low frequency range (Hons, 2008). One advantage of geophones over accelerometers is that they do not require electrical energy to give a signal output because they are transducers that transforms mechanical energy into electrical energy.
Fig. 4.37: Geophone. A coil is suspended from a frame by springs. When the frame moves, a relative motion appears between the magnet and the coil generating a voltage at its output.

Fig. 4.38: Generic geophone transference indicating its low sensitivity at low frequencies.
4.13 Acoustic Transducers

4.13.1 Introduction

Many devices based on electromagnetic wave propagation have been successfully developed with the purpose of studying environmental problems in the atmosphere. However the attenuation of the electromagnetic waves limits their underwater operation and hinders their underground application also. For these reasons, acoustic waves are employed instead of electromagnetic waves. Such are the cases of underwater communication between divers or underwater control of remote operated vehicles. Also, acoustic wave propagation is frequently employed to study the seabed. This is why acoustic transducers are so important in oceanography, hydraulics, geology and geophysics. This key role of the acoustic transducers in environmental research suggests developing this subject with some detail.

This topic dedicated to acoustic transducers will include both acoustic sensors and acoustic generators. They have several characteristics in common, and in many cases these transducers work in both ways; sending (generators) and receiving (sensors) acoustic signals.

According to their working principle, three groups of acoustic transducers should be considered; one based on Faraday’s law, another based on the piezoelectric effect and a third one based on the magnetostrictive effect. The first two groups are the most widespread and used.

Speakers and microphones used in audio to play and record voice and music are acoustic transducers designed to work in air. A magnet and a coil of wire are employed to transfer mechanical energy into electrical energy and vice versa. They are based on Faraday’s law and the Lorentz force, and are known as inductive transducers.

Most of the attention of this topic will be paid on piezoelectric transducers which are used in many applications related to environmental studies. Some sound generators based in the magnetostrictive effect have particular features and they will also be briefly described.

4.13.2 Inductive Transducers

Figure 4.39 depicts a generic bidirectional inductive transducer. It can work either as a sensor or a generator, as a microphone or as a loudspeaker. It has an element whose function is to transform wave pressure into motion or motion into pressure waves; it is a flexible component which is called diaphragm in microphones and cone in loudspeakers. The edge of this element is fixed to the frame but pressure waves can move it back and forth at the center, where a coil of wire is wound. A magnet is attached to the frame such that the coil is located in its magnetic field. Two electrical wires are connected to the ends of the coil.
4.13.2.1 Working as Sensor

When the transducer of Figure 4.39 works as a sensor or microphone, the mechanical energy from the pressure wave moves the diaphragm and is converted into electrical energy by the coil placed in the magnetic field. Thus, the input energy to the transducer is the sound and the output signal is a voltage between the two electrical wires.

This device is also known as constant-velocity microphone or moving coil microphone, and is based on Eq. (3.23) (Faraday’s law) which is applied to the microphone example and shown in Eq. (4.19),

\[ V_o = Blu \]  \hspace{1cm} (4.19)

where \( V_o \) is the output voltage (emf \( \varepsilon \)), \( B \) the magnetic field, \( l \) the conductor length of the coil and \( u \) the coil velocity.

---

![Diagram of a generic bidirectional inductive transducer](image-url)

**Fig. 4.39:** A generic bidirectional inductive transducer

4.13.2.2 Working as Generator

In an inductive generator an electrical energy is transformed into a pressure wave. In Figure 4.39 an alternating voltage is connected to the coil by the two electrical wires and current flows through the coil placed in a magnetic field. As a result a magnetic
force appears on the coil (Lorentz force) as shown in Eq. (3.28). This force displaces the center of the cone, thus generating a pressure wave. This is the way a loudspeaker works; if the current fluctuates at a given frequency, say 1000 Hz, a person will listen a sound of the same frequency. In general, the frequency range of these devices is between 20 to 20,000 Hz, so the frequency band of the voice is well reproduced.

### 4.13.3 Piezoelectric Transducers

Piezoelectric transducers are based on the piezoelectric effect (Section 3.8)) and are widely used as part of many instruments employed in environmental sciences. They are used to transform electric energy into mechanical energy and vice versa. Usually they work in air and water, but they have to be specifically designed for the environment where they will be used due to the different properties of media. Air is largely compressible and sound propagates in it at about 342 m/s, whereas water is very little compressible and sound travels in it at about 1480 m/s. Then the transference in each environment is different. It is said that the transducer “sees” different mechanical impedances in air than in water. Also, electrical parts in underwater sensors must be sealed from the medium, and the sensor’s housing has to be designed so as to support greater pressures than in air.

Figure 4.40 represents a piezoelectric disc with two electrodes; \( P \) is its polar axis and \( V \) the applied voltage.

![Piezoelectric disc with electrodes](image)

**Fig. 4.40:** Piezoelectric disc with electrodes. Voltage on the electrodes deforms the ceramic. Stress on the ceramic generates voltage between electrodes.
When the switch \((S)\) is closed, the piezoelectric disc changes shape from the solid line to the dashed line. As it was introduced in Section (3.8), a DC voltage applied to the bar will deform it constantly, whereas an AC electrical signal will make the disk vibrate at the frequency of the electrical signal. Conversely, a mechanical vibration applied to the material generates an electric alternating voltage of the same frequency than the vibration.

The electrical energy deforms the piezoelectric disc which, in turn, produces a mechanical energy proportional to the change in its length \(\Delta L = L_1 - L_2\). Therefore, with the purpose of increasing the converted energy the length variation must be maximized. With this purpose, several discs are usually stacked, alternating the direction of their polar axes to reduce the number of electrodes; electrical connections and electrodes \((E)\), are shown in Figure 4.41.

Therefore, the active vibrating element of some transducers consists of a stack of individual piezoelectric crystals. They can have shapes different to a disc but they are stacked so as to increase the energy transferred. A frequent piezoelectric transducer used in the sea is made of a stack of piezoelectric rings. They have the advantage that a bolt can be threaded through the hole to fix the stack.

\[ \begin{array}{c}
\text{- E} \\
\text{P} \\
\text{P} \\
\text{P} \\
\text{P} \\
\text{P} \\
\text{P} \\
\text{E} \\
\end{array} \]

Fig. 4.41: Piezoelectric ceramic discs are stacked to increase the transducer power. Their polar axe directions are alternated to reduce the number of electrodes. The symbols \(-E\) and \(+E\) should be considered as instantaneous values; half a cycle later the polarities are exactly the opposite.

### 4.13.4 Piezoelectric Sound Generators

A generic sound generator for use in the sea is depicted in Figure 4.42. It comprises three parts: the tail mass, the head mass and the active elements. Because these generators work underwater most of their parts are incorporated in a waterproof housing (not
shown). The active elements are piezoelectric rings threaded by a stud bolt. The bolt is screwed at both ends into the tail mass and the head mass compressing the rings. Then, when the piezoelectric elements are excited by an alternating voltage ($V$) they vibrate. Because the tail mass is big it has a great inertia and remains at rest. The head mass which is much lighter than the tail mass is pushed away from the tail mass producing a pressure wave in the water. This pressure wave will have the shape of the voltage source waveform (sinusoidal, pulsed, etc.).

![Piezoelectric active elements](image)

Fig. 4.42: The sound generator is composed of a tail mass, a head mass and the piezoelectric active elements. These elements move the head with respect to the tail generating a pressure wave.

Table 4.4 shows the technical characteristics of two quite different transducers for use in air. Typical applications of piezoelectric transducers in environmental sciences for use in air are: tide gauges, ultrasonic anemometers, acoustic radars (for measuring the atmospheric boundary layer and the wind speed) and geophones. Transducers for water are used in acoustic current meters, echo sounders, side scan sonars, water level meters, ocean and shallow water survey, commercial fishing equipment, instruments for measuring sediment concentration, underwater voice and data communications, submersible positioning, obstacle avoidance and hydrophones, among others.

Sound transducers for higher frequencies (in the order of megahertz) have the same principle but a different embodiment (Fig. 4.43). The same active element is generally used to emit and receive sound. The backing material is usually a high density material that absorbs part of the energy radiated backwards. The acoustic properties of this material define the spatial resolution (beamwidth) and the output signal amplitude or the receiving sensitivity.
In those cases where it is not required a great amount of energy, a single piezoelectric element may be used (stack of elements is not necessary). As a sound generator, the piezoelectric crystal is energized by a voltage source through the connector, the electrical wires and the electrodes. Signals follow the inverse path when the device works as a receiver.

Table 4.4: Piezoelectric transducers in air

<table>
<thead>
<tr>
<th>Operating Frequency: 300 kHz</th>
<th>Operating Frequency: 30 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical Sensing Range: 5 cm to 50 cm</td>
<td>Typical Sensing Range: 80 cm to 25 m</td>
</tr>
<tr>
<td>Beamwidth: 10° ± 2°</td>
<td>Beamwidth: 12° ± 2°</td>
</tr>
<tr>
<td>Diameter = 12 mm</td>
<td>Diameter = 106 mm</td>
</tr>
<tr>
<td>Length = 10 mm</td>
<td>Length = 141 mm</td>
</tr>
</tbody>
</table>

The protective plate protects the transducer from the environment. For underwater transducers it has also the purpose of serving as an acoustic transformer between the high acoustic impedance of the active element and the water.
4.13.5 Hydrophones

A hydrophone (Figure 4.44) is a sensor that works submerged in a liquid, generally water. It receives sound energy (pressure) and produces an electrical output. It works like an underwater microphone. A hydrophone is a listening passive device usually composed of several piezoelectric rings threaded by a stud bolt screwed into both end heads (Naval Underwater Systems Center, 1990).

Fig. 4.44: Hydrophone. A stack of piezoelectric rings connected to sum up the voltage generated by each one. The electrical output is conducted through the submersible cable to the amplifier. The set is sealed by a rubber cover.

Ring elements have metallic electrodes which are connected alternating the direction of the polar axis (to reduce the number of electrodes) as illustrated in Figure 4.41. A cover of rubber, which is shown cut in the figure, protects the positive and negative electrodes, the internal connections and the ceramic rings from water. The wires from the electrodes are connected to a submersible cable which leads the signal to an amplifier. The schematic presented here is one of the several embodiments that a hydrophone may have.

4.13.6 Magnetostrictive Transducers

Magnetostrictive transducers use the magnetostrictive effect (Section (3.9)) of a material to convert magnetic energy into mechanical energy and vice-versa. A kind of magnetostrictive transducer for low power applications may be built arranging in
parallel a large number of plates made of a magnetostrictive material. A coil of wire is placed around the plate’s stack (Fig. 4.45). When an alternating current \((I)\) is forced through the coil, a magnetic field causes the magnetostrictive material to change its length, resulting in a mechanical vibration.

**Fig. 4.45:** A magnetostrictive transducer formed by plates of magnetostrictive material and a coil. Voltage power supply generates the magnetic field which makes the plates vibrate and change their length as shown by the arrows.

The optimal operation frequency in magnetostrictive transducers depends on the length of the transducer; higher frequencies require shorter lengths. Physical size limitations restrict the operating frequencies much below those reached with piezoelectric sensors (http://www.ctgclean.com).

### 4.14 Rotation Sensors

In many measuring systems it is required to know the angle of rotation of a shaft, as in the case of an anemometer’s direction sensor (vane) or the angular speed as in the cups and propeller of wind speed sensors.

A simple way to measure both parameters is to couple a certain kind of piece to the shaft. This piece is just used to amplify the mechanical information, facilitating the mounting of the required measuring sensors. The shape and material of the piece will depend on the operating principle selected for the sensor. There are several classic operating principles based on optical and electrical effects.
4.14.1 Rotation Speed

We have described in Section (3.7.13) (Fig. 3.22) the way in which electrical energy can be generated from mechanical energy. A magnet was coupled to a rotating shaft and coils placed in the varying magnetic field to generate an electromotive force (voltage). A pulse is generated when one of the magnetic poles passes close to the coil. Then, counting the amount of pulses produced by the coils in a certain time, the number of turns of the shaft may be estimated.

Also, it was stated that the amplitude of the generated voltage is proportional to the time rate of change of the magnetic flux, which is proportional to the rotation speed of the magnet. The average voltage output is thus proportional to the average rotation speed of the shaft. Then coupling a voltage generator to the shaft is another way used to know its rotation speed.

A different electrical device for measuring the above mentioned magnitude is depicted in Figure 4.46, where a voltage is applied to the coil terminals producing a magnetic field around the coil.

![Diagram](image)

**Fig. 4.46:** The voltage applied to the coil produces a magnetic field that is disturbed by the rotating protrusions. The alteration of the magnetic field produces current changes in the coil. Counting the current changes permits protrusions passes and shaft rotation to be determined.
As explained above for inductive sensors (Section (4.4.2)), an electrical conductive piece close to the coil would disturb the produced field, modifying the current flowing through the coil. A conductive rotary piece with protrusions is then coupled to the shaft. The rotation of the shaft will produce a pulse-shaped disturbance of the magnetic field so that a pulsed perturbation of the current will be generated in the coil. This perturbation is amplified and analogically processed by an electronic circuit to generate a clear pulsed signal which in turn is counted over a known period of time. The amount of pulses counted is proportional to the average speed of the shaft.

A different operating principle consists in coupling to the shaft a disc with small magnets attached to it. A variable magnetic field in the vicinity of the disc is thus generated when the shaft rotates (Fig. 4.47). This changing field can be detected by a magnetic field detector which could be a Hall effect switch (Section 3.7.18), a reed switch (Section 3.7.19) or a coil; as explained in Section (3.7.12), if a magnet moves in the vicinity of a coil a voltage will be produced at the coil terminals.

When the magnets pass close to the magnetic field detector, the detector will perceive changes in the magnetic field and an electronic circuit associated with it will produce pulses. Again, counting pulses over a fixed period of time gives the average rotating speed.

![Diagram of rotation sensors](image_url)

**Fig. 4.47:** A disc is attached to the shaft and magnets are placed on the disc. The magnetic field detector may by a coil, a Hall effect switch or a reed switch.

In the case of the coil, some mechanical energy from the shaft is taken by the coil to produce the electrical signal. It can be thought that the induced current circulating by the coil becomes an electromagnet that attracts the passing magnet tending to stop the disc. Then it is said that the shaft suffers a “mechanical charge” that opposes its rotation and it appears a braking torque on the shaft. This torque is of electrical origin.
and must be differentiated from the friction torque. In those methods where a sensing coil is used, a change in the coil current is produced by the change in the magnetic field; this means that an electrical energy is generated by means of the rotating shaft through the measuring system, and also some electrical braking torque appears. No electrical energy can be created without some mechanical work.

Coils are thus not the most adequate device for detecting rotation when the mechanical energy available is low because they tend to stop the rotation of the shaft. This effect is observed as a threshold in the transference of the instrument, which is a certain inability of the device to rotate when the velocity is low.

Another way of measuring the speed of a rotating shaft that does not have the above problem is by using a slotted disk to interrupt the optical path between an emitter and a receiver (Fig. 4.48). The emitter may be a LED and the receiver a photo detector. In the example of Figure 4.48 the disc lets the light beam pass twice per turn, generating two narrow pulses in one complete rotation. These devices are known as encoders, and today’s technology permits them to be manufactured so as to produce thousands of pulses per turn by adding a large number of slots.

**Fig. 4.48:** A chopper disc attached to the shaft interrupts a light beam. The amount of pulses received in a known period is counted and it is proportional to the shaft speed.

It is worth noting that there is also a threshold in the transference of the instrument in the optical case, but it is only due to the friction and inertia of the mechanical parts.
4.14.2 Rotation Angle

The rotation angle of a shaft may be known by an optical system. It consists in attaching a coded disc to the shaft. The disc should have transparent and opaque zones which let light pass, or stop it. The shape and distribution of these zones are appropriately selected so that the detection of light through the disc permits the disc position to be determined. Figure 4.49 is the top view of one of such discs. Each of the three small squares represents a pair of optical transmitter and receiver as seen from above.

![Disc Diagram](image)

**Fig. 4.49:** A three bit coded disc. The squares are the light beams perpendicular to the disc plane. The binary code for each position is illustrated below each disc. For example the bottom left disc shows 0-1-0 which indicates that the first and the last beam are in dark zones, then, they are interrupted and the output at the receiver is “0”. The central beam passes by the disk and is detected by the receiver giving as a result a digit “1”.

Each light beam is perpendicular to the disc plane, and the optical pairs are disposed as was shown in Figure 4.48. In this example, three circular bands have been coded with transparent and opaque zones. The central zone is all transparent, so it does
not contribute to the angular detection. Each receiver has only two possible states: the disc let the light pass, and its output is high, or the disc obstructs the light beam and its output is low. Because of this restricted dual state capability it is an inherently digital system. The detection of light through the disc is illustrated below the discs, binary-coded for each position. Each optical pair corresponds to a bit of the digital word and with three bits it is possible to determine $2^3 = 8$ positions, then the resolution of this example is $45^\circ$.

The angular resolution of this rotation sensor in increased by accommodating more circular bands on the discs and placing more optical pairs, which requires larger discs. At present, 12 bit angular resolution ($360^\circ/2^{12} = 0.088^\circ$) is achieved with discs whose diameters are about 75 mm. Usually, sensor digital outputs are acquired by microprocessors and presented on a display or stored in memory. This kind of sensor is very reliable, has long life and requires very low torque to rotate the disc, but the optical emitters consume non-negligible amounts of electrical power.

A more simple rotation angle detector may be implemented with a circular potentiometer (Fig. 4.50). Both end terminals of the potentiometer must be connected to an electrical power supply, and the shaft whose angular displacement has to be measured, must be coupled to the sliding terminal. The position of this terminal over the potentiometer resistance defines the output voltage. Thus, voltage is directly related to the angular position. The transference curve of this sensor is liner and the constant transference is expressed in units of voltage per degree. For example, for $0^\circ$ the output will be $0 \, \text{V}$, for $180^\circ$ will be half the power supply voltage ($V/2$) and for $360^\circ$ the output will be $V$.

![Fig. 4.50: A circular potentiometer with the sliding terminal attached to the shaft and the end terminals to the voltage supply. Voltage at the sliding terminal is proportional to the shaft angle.](image-url)
In general, this kind of potentiometer has a small gap between the two end terminals which precludes measuring over the entire 360°. Furthermore, because there is a friction between the sliding terminal and the body of the potentiometer, the shaft needs to produce a certain torque to move the sliding terminal. For example, if the shaft is coupled to a vane to measure wind direction, low winds could not have enough energy to displace the sliding terminal. Fortunately, in general, low winds are of little interest. Also, for continuous moving shafts, the friction could wear the potentiometer adding noise to the output voltage.

Sometimes, it is desired to average the output of the angular sensor which is vibrating about the real direction when, for example, it is desired to obtain an average wind direction or to decrease random noise. Using potentiometers, the averages have to be performed taking some precautions, because when the sliding terminal is jumping between the two end terminals, say between 0° and 360°, a simple average will give as a result 180° which is almost the opposite to the real direction.

4.15 Concluding Remarks

As was noted in Section (1.7), the sensors we are interested in are those with electrical outputs. Therefore, to begin the description of these sensors it was needed to introduce the principles on which they are based on. Thus, it was unavoidable an introduction to electrical and electronic circuits to familiarize instrument users with general concepts needed to understand their working principles.

Subsequently, it was necessary to study the most elementary sensors which are based on changes produced by measurands on resistances, capacitances and inductances. Once these concepts were explained, other more realistic sensors based on these principles were studied: sensors to measure temperature, force, pressure, humidity, fluid conductivity, acceleration, sound, vibration, distance, velocity, rotation, position, etc.

Also, to help understanding automatic measurement systems and oscillator principles, it was needed to introduce some simple inductive actuators. Whenever time must be measured, or some kind of periodic waveform must be generated, oscillators are necessary. Then a conceptual approach to them was developed.

The purpose of the sensor description made here was to introduce readers to the essential nature of sensors with electrical output. It is expected that understanding these basic concepts will help readers to understand any other new sensor that they would find in the future.

Even when the sensors described are used in many fields of science and technology, it must be emphasized the role they have as parts of instruments used in Environmental Sciences. In the following paragraphs there is a summary relating these sensors with those instruments in which they could be used. Finally, in Table 4.5 there is a synopsis of sensor specifications and their applications.
Temperature sensors whose characteristics were outlined in Table 4.1 are extensively used in Environmental Sciences to record the temperature of soil, air and water. Also more sophisticated instruments such as infrared thermometers and infrared thermography are used to assess plant canopy temperature; studies on land-atmosphere interaction, surface energy balance; human and animal body temperature measurements, etc. They are also used as part of radiometers and of groundwater flowmeters.

Pressure sensors are also of massive use; their specifications and application depend on the technologies used for manufacturing them. In our field of interest, they are used to measure groundwater and river levels, wind waves, tsunamis, atmospheric pressure, pressure of drilling systems (geology), pressure on the wings of birds and even insects (http://environmentalresearchweb.org/cws/article/news/50265).

Humidity sensors are used in instruments to record humidity in soils, air, woods and grains (wheat, barley, corn, soil bean, etc.); they are also applied in irrigation studies and research of evaporation processes and soil–atmosphere interaction.

Electrical conductivity of liquids is measured to know the amount of dissolved salts, to detect the trajectory of tracers in groundwater, to detect the presence of contaminants, to evaluate aquifer risk under exploitation, etc.

Accelerometers are used in studying the dynamics of animals and human beings; they are deployed in ocean buoys to measure waves, also in robots to study hard to reach places such as deep oceanic waters or the surface of the moon. They are also used to measure speed and displacement in underwater vehicles for oceanographic research.

Geophones are devices used to measure small earth vibrations. They are used to study different layers of soils and to estimate their components, and are required in studies to detect the presence of minerals, water or oil.

Acoustic transducers are present in countless applications either as acoustic receivers or acoustic generators. They are used to measure water level, waves, the speed of particles in air or water, the speed of sound, wind speed, thickness of tissues, ocean currents, and water velocity in rivers and in the laboratory. They are also used in underwater vehicles positioning systems, flowmeters, underwater communication equipments, etc.

Rotation Speed and Angle sensors are used to measure heave, pitch and roll of sea buoys or vehicles, in instruments such as anemometers or groundwater velocimeters.

Strain gauges are used, for example, to measure dynamic forces exerted by living beings and wind force on trees or crops (Cleugh et al., 1998).
**Table 4.5: Synopsis of sensors**

<table>
<thead>
<tr>
<th>Measurand</th>
<th>Approximate full scale error</th>
<th>Field of applications in Environmental Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pressure (semiconductor)</td>
<td>±1% - ±10%</td>
<td>Low accuracy and low cost applications.</td>
</tr>
<tr>
<td>Pressure (strain gauges)</td>
<td>±0.1% - ±1%</td>
<td>Water table, tides, waves, atmospheric pressure, pressure of drilling systems.</td>
</tr>
<tr>
<td>Pressure (quartz crystal)</td>
<td>0.0001% - 0.01%</td>
<td>Precision water levels, tsunamis.</td>
</tr>
<tr>
<td>Humidity (conductive)</td>
<td>5% - 100%</td>
<td>Soil humidity.</td>
</tr>
<tr>
<td>Humidity (capacitive – full range)</td>
<td>±2% - ±5%</td>
<td>Air, gas.</td>
</tr>
<tr>
<td>Humidity (capacitive – limited range)</td>
<td>± 0.5%</td>
<td>Grains.</td>
</tr>
<tr>
<td>Electrical conductivity of fluids</td>
<td>±1% - ±5%</td>
<td>Dissolved salts, contaminants, and aquifer risk evaluation.</td>
</tr>
<tr>
<td>Acceleration</td>
<td>≈ 5%</td>
<td>Dynamics of living beings, waves, underwater vehicles.</td>
</tr>
<tr>
<td>Soil vibration</td>
<td>±2.5%</td>
<td>Geology, detection of minerals, water and oil.</td>
</tr>
<tr>
<td>Acoustic waves</td>
<td>Depends on the application</td>
<td>Water level, waves, thickness of tissues, underwater communications, air and water velocity profilers.</td>
</tr>
<tr>
<td>Rotation speed and angle</td>
<td>Depends on the application</td>
<td>Measure of heave, pitch and roll. Anemometers, winches.</td>
</tr>
<tr>
<td>Strain gauges</td>
<td>≈ 1%</td>
<td>Forces exerted by living beings, wind force on trees or crops.</td>
</tr>
</tbody>
</table>

**References**


http://www.x26.com/articles.html#

http://www.paroscientific.com/qtechnology.htm

http://www.ctgclean.com/technology-library/articles/magnetostrictive-versus-piezoelectric-transducers-for-power-ultrasonic-applications/