Chapter 3

FIGURES OF MERIT FOR SENSING

Reference textbook*:* J. Fraden – *"Handbook of modern sensors",* Springer (2016)

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

What is a sensor?

A sensor is often defined as a “device that receives and responds to a signal or stimulus”. This definition is broad. In fact, it is so broad that it covers almost everything from a human eye to a trigger in a pistol. This world is divided into natural and man-made objects.

The natural sensors, like those found in living organisms, usually respond with signals having **electrochemical characteristics** (i.e., the nerve fibers). In **man-made devices**, information is also transmitted and processed in **electrical form**, through the transport of electrons. Therefore, a narrower definition of a sensor, may be phrased as:

“A sensor is a device that receives a stimulus and responds with an electrical signal.”

In this class, are also included all the **sensors for environmental monitoring.**

The term stimulus represents the quantity, property, or condition that is received and converted into electrical signal, either voltage, current or charge.

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

What is a sensor?

Any **sensor is an energy converter**, as they all deal with energy transfer between the object of measurement to the sensor. The process of sensing is a particular case of **information transfer**, thus the sensor’s performance must be assessed only as part of a **data acquisition system.**

The term **sensor** and term **detector** are typically used as **synonyms**. However, detector is more often used to stress qualitative rather than quantitative nature of measurement.

The term **sensor is distinguished from transducer**. The latter is a converter of any one type of energy or property into another type of energy or property, whereas the former converts it into electrical signal. However, a sensor can incorporate different transducers.

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

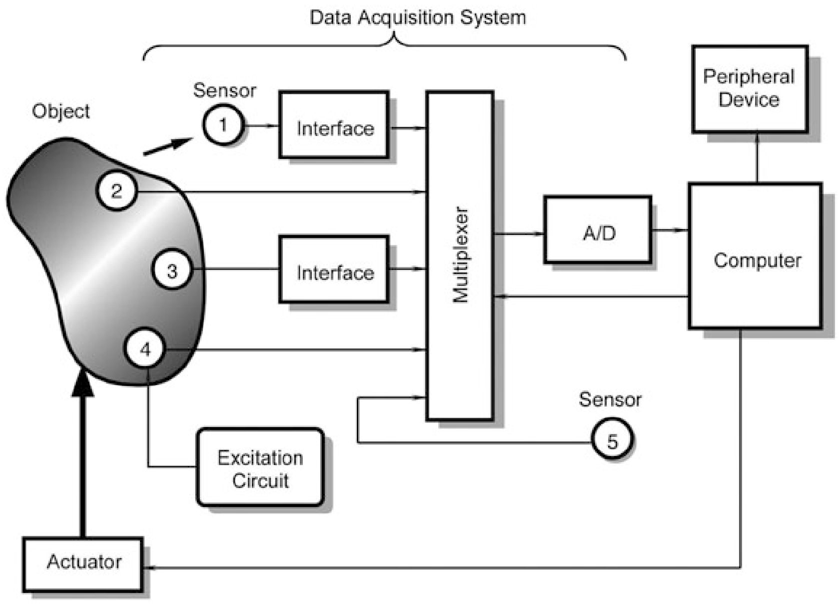
What is a sensor?

In general, **a sensor does not operate standalone**, but it is always part of a larger system that may incorporate other detectors, signal conditioners, memory devices, data recorders, and actuators.

Data can be collected from an object by several sensors.

Some of them are positioned directly on or inside the object. Other sensors perceive the object without a physical contact and are called noncontact sensor. Examples of such a sensor is a radiation detector and a TV camera.

Additional sensors can be used to monitor the internal conditions of the data acquisition system itself. Moreover, some sensors cannot be directly connected to standard electronic circuits because of the inappropriate output signal formats. They require the use of interface devices (signal conditioners) to produce a specific output format.



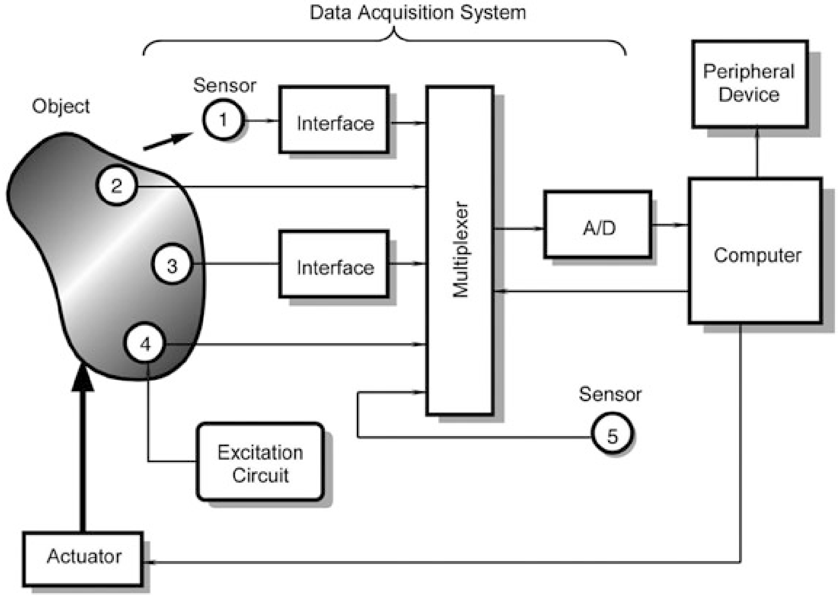
3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

What is a sensor?

Electrical signals from multiple sensors are fed into a **multiplexer (MUX)**, which is a switch or a gate.

Its function is to connect the sensors, one at a time, to an **analog-to-digital converter (A/D or ADC)** if a sensor produces an analog signal, or directly to a computer if a sensor produces signals in a digital format.

The computer controls a multiplexer and ADC for the appropriate timing. Also, it may send control signals to an actuator that acts on the object. Examples of the actuators are an electric motor, a solenoid, a relay, and a pneumatic valve.



3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

Sensor classification

Sensor classification schemes range from very simple to the complex. Depending on the classification purpose, different classification criteria may be selected. However, several **practical ways to look at sensors exist.**

- **Passive and active sensors**

  A **passive sensor** does not need any additional energy source. **It generates an electric signal in response to an external stimulus.** That is, the input stimulus energy is converted by the sensor into the output signal. Examples: thermocouples, photodiodes.

  An **active sensor** requires external power for their operation, which is called an excitation signal. That signal is modified (modulated) by the sensor to produce the output signal. The active sensors sometimes are called parametric because their own properties change in response to an external stimulus and these properties can be subsequently converted into electric signals. It can be stated that **a sensor's parameter modulates the excitation signal, and that modulation carries information of the measured value**. Examples: thermistors, resistive strain gauge.

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

Sensor classification

- **Absolute and relative sensors**

  An **absolute sensor** detects a stimulus in reference to an **absolute physical scale** that is independent on the measurement conditions. Examples of absolute sensors are thermistors, as their electrical resistance directly relates to the absolute temperature scale of Kelvin, or absolute pressure sensors, which produce signals in reference to an absolute zero on a pressure scale, i.e., the vacuum.

  A **relative sensor** produces a signal that **relates to some special case**. Examples of relative sensors are thermocouples as they produce an electric voltage that is function of a temperature gradient across the thermocouple wires. Thus, a thermocouple output signal cannot be related to any particular temperature without referencing to a selected baseline. In the same way, A relative pressure sensor produces signal with respect to a selected baseline that is not zero pressure, for example, to the atmospheric pressure

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

Sensor classification

- **Sensor's characteristics**

  Another way to look at a sensor is to consider some of its properties that may be of a specific interest. These properties can be grouped as:

  - **sensor specifications** (sensitivity, range, stability, resolution, accuracy, selectivity, speed of response etc.);
  - **sensing element material** (inorganic, organic, conductor, insulator, semiconductor, liquid, gas or plasma etc.);
  - **conversion phenomena**, which are mainly physical (thermoelectric, photoelectric, magnetoelectric, electromagnetic, thermoelastic, etc.) or chemical (chemical and physical transformation, electrochemical processes, etc.)
  - **stimuli** (acoustic, electric, magnetic, optical, mechanical, thermal, etc.)

3.1 DATA ACQUISITION AND SENSOR CLASSIFICATION

Units of Measurements

Any physical or chemical measure must be related to a unit of measurements. The only exception are adimensional parameters and normalized scales.

The base measurement system is known as **SI** which stands for French “Le Systéme International d’Unités”, which sets the **basic units**. All other physical quantities are **derivatives of these base units**.

Quantity	Name	Symbol
Length	meter	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candle	cd

3.2 TRANSFER FUNCTION

What is a transfer function?

Most of stimuli provided to a sensors are not electrical. Therefore, from its input to the output a sensor may perform several **signal conversion** steps before it produces and outputs an electrical signal.

For example, in fiber optic pressure sensor, the applied pressure results in strain in the fiber which in turn causes deflection in its refractive index, which in turn changes the optical transmission and modulates the photon density and finally the photon flux is detected by a photodiode and converted into electric current.

How can we model the input-output relation? How can we employ this relation to determine an unknown input stimulus from the sensor's electric output?

3.2 TRANSFER FUNCTION

What is a transfer function?

An ideal or theoretical **input–output (stimulus–response) relationship** exists for every sensor. For an ideal sensor, the output would always represent the true value of the stimulus. This ideal input–output relationship may be expressed in the form of a table of values, graph, mathematical formula, or as a solution of a mathematical equation.

If the input–output function is time invariant, it is called ***transfer function***.

A transfer function represents a relation between the input stimulus ***s*** and the electrical signal ***E*** produced by the sensor at its output. This relation can be written as $\boldsymbol{E = f(s)}$.

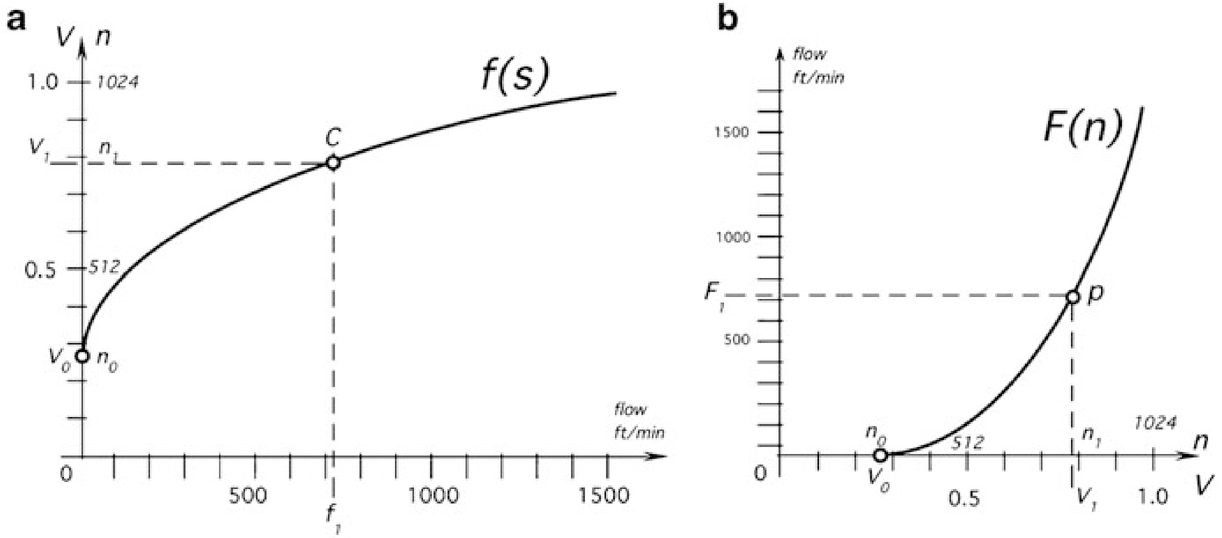
Normally, the stimulus is unknown while the output signal is measured and thus becomes known. The value of *E* that becomes known during measurement is a number (voltage, current, digital count, etc.) that **represents stimulus *s***.

3.2 TRANSFER FUNCTION

What is a transfer function?

In real life, any sensor is related to a **measuring system**. One of the functions of the system is to “break the code *E*” and **infer the unknown value of s from the measured value of *E***. Thus, the measurement system shall employ an inverse transfer function $\boldsymbol{s = f^{-1}(E) = F(E)}$, to compute value of the stimulus *s*. It is usually desirable to determine a transfer function not just of a sensor alone, but rather of a system comprising the sensor and its interface circuit.

Graphically, the inverse function can be obtained by a mirror reflection with respect to the bisector of the right angle formed by x and y-axes.



3.2 TRANSFER FUNCTION

Mathematical models of transfer functions

Preferably, a physical or chemical law that forms a basis for the sensor's operation should be known. If such a law can be expressed in form of a mathematical formula, often it can be used for calculating the sensor's inverse transfer function by inverting the formula and computing the unknown value of s from the measured output *E*.

For example, a linear resistive potentiometer is used for sensing displacement *d*, representing the stimulus *s*. The Ohm's law can be applied for computing the transfer function. In this case, the electric output *E* is the measured voltage *V* while the inverse transfer function is given as:

$$d = F(E) = \frac{D}{V_0} V \tag{3.1}$$

where V_0 is the reference voltage and *D* is the maximum displacement (full scale), both being the constants. By using this function, we can compute displacement *d* from the measured voltage *V*.

In practice, **readily solvable formulas for many transfer functions, especially for complex sensors, does not exist** and various approximations of the direct and inverse transfer functions are used.

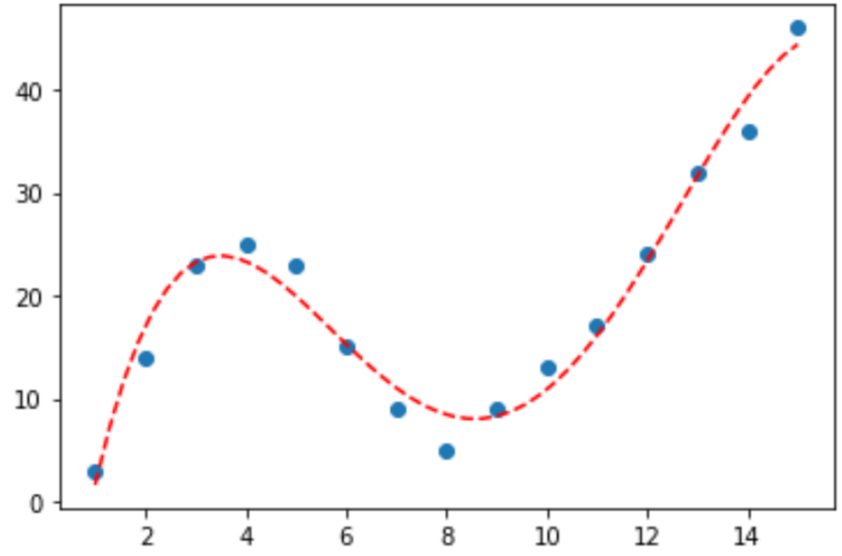
3.2 TRANSFER FUNCTION

Functional approximations

Approximation is a selection of a suitable mathematical expression that can fit the experimental data as close as possible. **The act of approximation can be seen as a curve fitting of the experimentally observed values into the approximating function**.

The approximating function should be simple enough for ease of computation and inversion and other mathematical treatments, for example, for **computing a derivative to find the sensor's sensitivity**. The selection of such a function requires some mathematical experience.

Initially, one should check if one of the basic functions can fit the data and if not, then resort to a more general curve-fitting technique.



3.2 TRANSFER FUNCTION

Functional approximations

The simplest model of a transfer function is linear. It is described by the following equation:

$$E = A + Bs \tag{3.2}$$

The intercept *A is* the output signal *E* at zero input signal *s=0*. **The slope of the line is *B*, and it is also called sensitivity** since the larger this coefficient the greater the stimulus influence. The slope *B* is a tangent of the angle *α.*

Eq. (3.2) assumes that the transfer function passes through zero value of the input stimulus *s*. **In many practical cases it is just difficult or impossible to test a sensor at a zero input**. For example, a temperature sensor used on a Kelvin scale cannot be tested at the absolute zero (-273.15 °C).

Thus, in many linear or quasilinear sensors it may be desirable to reference the sensor not to the zero input but rather to some more practical **input reference value s_0**.

3.2 TRANSFER FUNCTION

Functional approximations

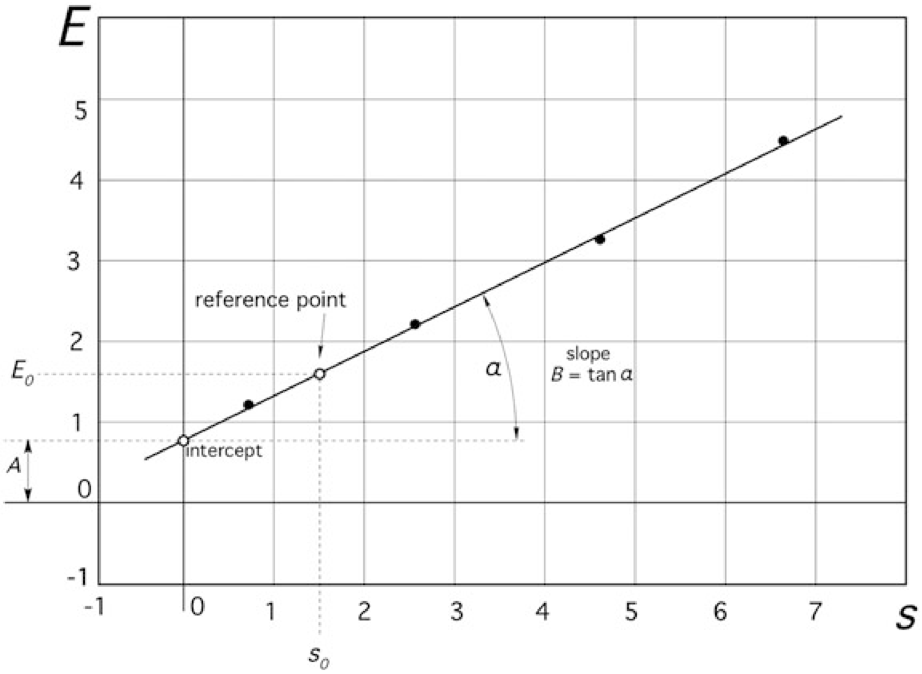
If the sensor response is E_0 for some known input stimulus s_0, Eq. (3.2) can be rewritten as:

$$E = E_0 + B(s - s_0) \qquad (3.3)$$

The **reference point has coordinates s_0 and E_0**.

The **inverse linear transfer function** for computing the input stimulus from the output E is:

$$s = \frac{(E - E_0)}{B} + s_0 \qquad (3.4)$$



3.2 TRANSFER FUNCTION

Functional approximations

In the real world, **nonlinearities are often present especially for a broad input range of the stimuli**. Thus, Eqs. (3.2) and (3.3) represent just a linear approximation of a nonlinear sensor's response, where a nonlinearity can be ignored for the practical purposes or in a specific operating range.

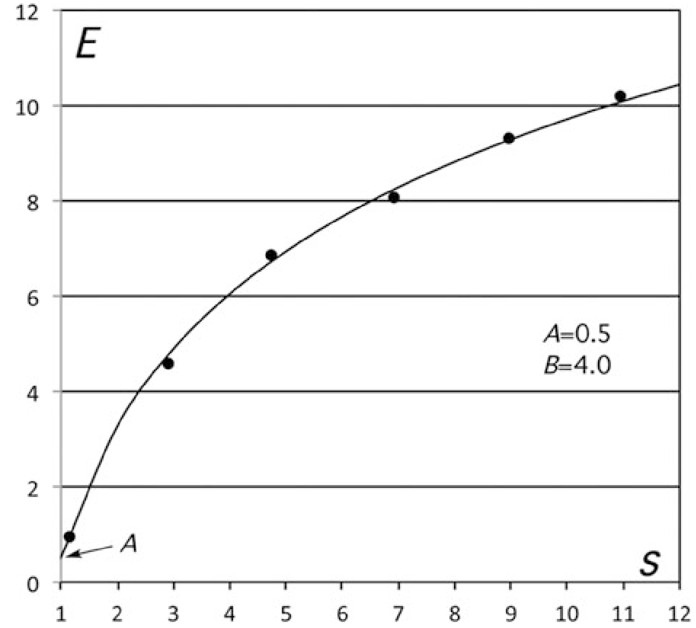
A nonlinear transfer function can be approximated by a **nonlinear mathematical function**, as logarithmic, exponential and power function.

The **logarithmic approximation function** and the corresponding inverse function (which is exponential) are respectively:

$$E = A + B\ \ln(s) \qquad (3.5)$$

$$s = e^{\frac{E-A}{B}} \qquad (3.6)$$

where A and B are the fixed parameters.



3.2 TRANSFER FUNCTION

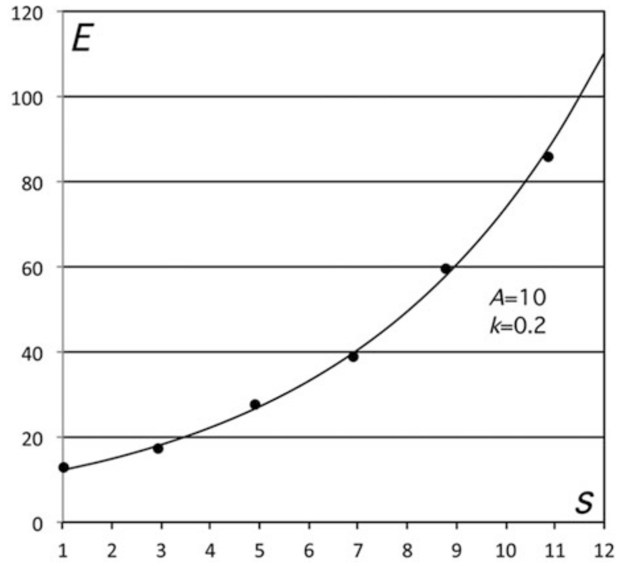
Functional approximations

The **exponential function** and its inverse (which is logarithmic) are given by:

$$E = Ae^{ks} \tag{3.7}$$

$$s = \frac{1}{k}\ln\left(\frac{E}{A}\right) \tag{3.8}$$

Where A and k are the fixed parameters.

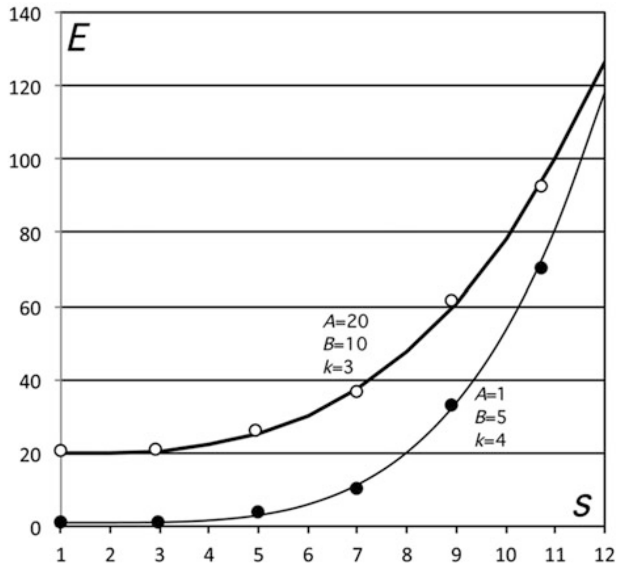


The **power function** and its inverse are expressed as:

$$E = A + Bs^{k} \tag{3.7}$$

$$s = \sqrt[k]{\frac{E-A}{B}} \tag{3.8}$$

Where A and B are the fixed parameters, while k is the power factor.



3.2 TRANSFER FUNCTION

Functional approximations

For a nonlinear transfer function, sensitivity is not a fixed number, as would be the case in a linear transfer function.

A nonlinear transfer function exhibits different sensitivities at different points in intervals of stimuli. In the case of nonlinear transfer functions, sensitivity is defined as a first derivative of the transfer function at the particular stimulus s_i:

$$b_i(s_i) = \frac{dE(s_i)}{ds} = \frac{\Delta E_i}{\Delta s_i} \tag{3.7}$$

where, Δs_i is a small increment of the input stimulus and ΔE_i is the corresponding change in the sensor output E.

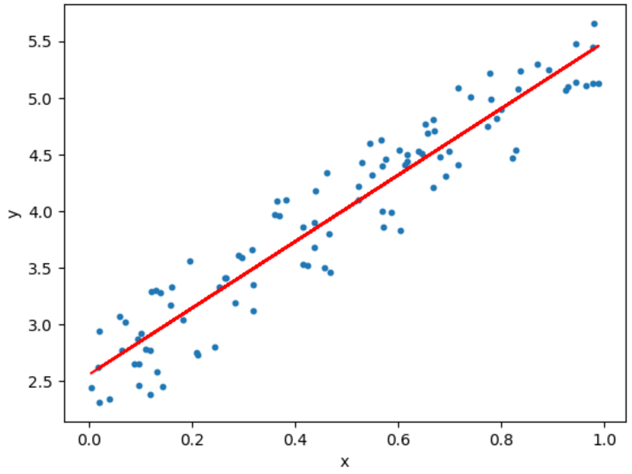
3.2 TRANSFER FUNCTION

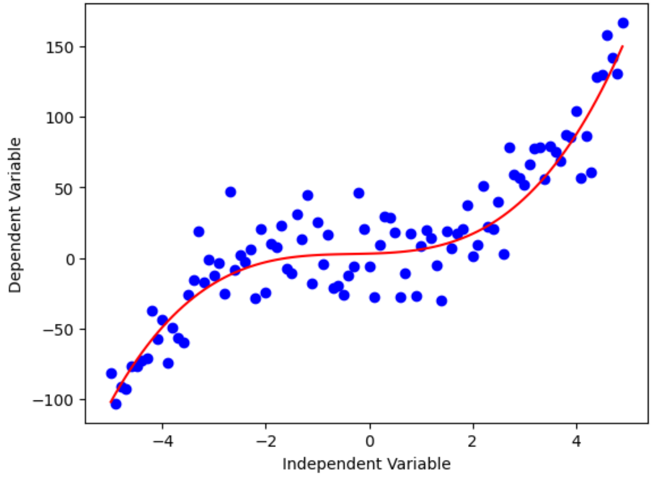
Functional approximations

The **measurements** of the input stimuli during calibration is always characterized by a certain **degree of accuracy and precision**. To cope with random errors in the calibration process, regression methods are employed to find the function parameters.

The most used **linear regression method is based on the ordinary least squares**.

Nonlinear regression methods rely on different approaches, such as **data linearization**, **polynomial** approximation, and **non-linear least squares** algorithm (e.g., Levenberg-Marquardt algorithm).





3.2 TRANSFER FUNCTION

Multidimensional Transfer Functions

A sensor transfer function may depend on more than one input variable. That is, the sensor's output may be a function of several stimuli.

- One example is a humidity sensor whose output depends on two input variables: relative humidity and temperature.
- Another example is the transfer function of a thermal radiation (infrared) sensor. This function has two arguments, namely two temperatures: T_b, the absolute temperature of an object of measurement and T_s, the absolute temperature of the sensing element. Thus, the sensor's output voltage V is proportional to a difference of the fourth-order parabolas:

$$V = G\left(T_b^4 - T_s^4\right) \tag{3.8}$$

where G is a constant.

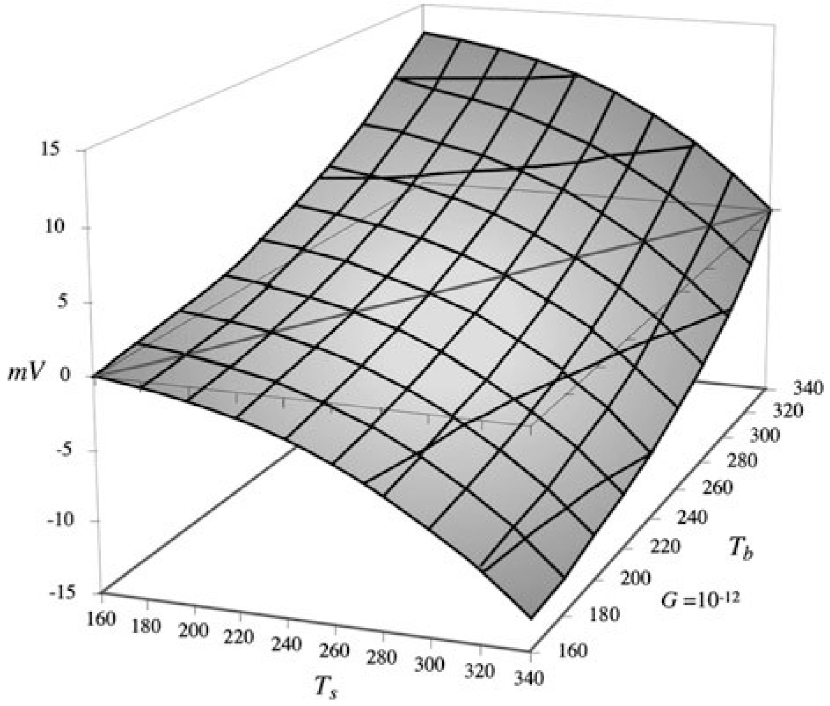
3.2 TRANSFER FUNCTION

Multidimensional Transfer Functions

The relationship between the object’s temperature T_B and the output voltage V is not only nonlinear but also in a nonlinear way depends on the sensing element surface temperature T_s.

To determine the sensitivity of the sensor with respect to the object’s temperature, a partial derivative will be calculated as:

$$b = \left.\frac{dV}{dT_b}\right|_{T_s} = 4GT_b^3 \tag{3.9}$$



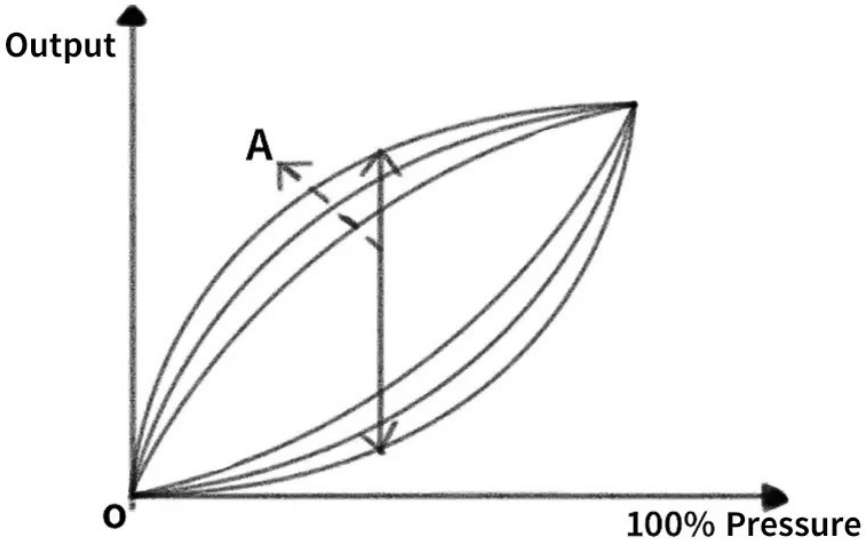
3.3 SENSOR CHARACTERISTICS

Sensor characteristics

When selecting a sensor, the first thing to do is to **outline requirements for the targeted application**. When knowing what is needed, one is ready to evaluate what is available. The evaluation starts by studying the sensor’s data sheet that specifies all its essential characteristics. The task then is to match the requirements to availability.

The main sensor characteristics include:

- Dynamic range
- Accuracy
- Hysteresis
- Saturation
- Repeatability
- Resolution
- Dynamic response



3.3 SENSOR CHARACTERISTICS

Dynamic range

A **dynamic range of stimuli** which may be converted by a sensor is called **a span or an input full scale (FS).** It represents the highest possible input value that can be applied to the sensor **without causing an unacceptably large inaccuracy**.

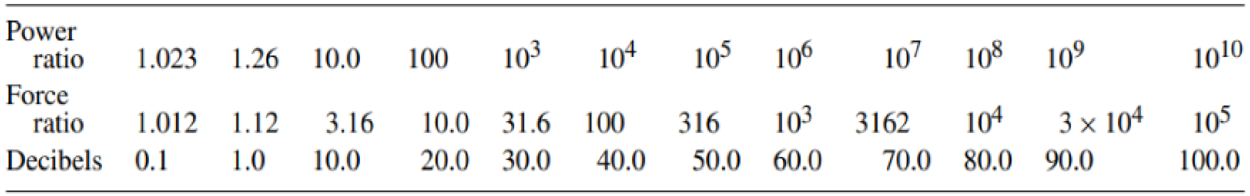
For the sensors with a very broad and nonlinear response characteristic, a dynamic range of the input stimuli is often expressed in **decibels**, which is a logarithmic measure of ratios of either power or force (voltage). In the case of ratios between powers, 1 dB is defined as:

$$1\ dB = 10 \log\left(\frac{P_2}{P_1}\right) \tag{3.10}$$

Similarly, in the case of ratios between voltages:

$$1\ dB = 20 \log\left(\frac{V_2}{V_1}\right) \tag{3.11}$$

Decibels do not measure absolute values, but a **ratio of values** only. Being a nonlinear scale, it may represent low-level signals with high resolution while compressing the high-level numbers.



Power ratio	1.023	1.26	10.0	100	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}
Force ratio	1.012	1.12	3.16	10.0	31.6	100	316	10^3	3162	10^4	3×10^4	10^5
Decibels	0.1	1.0	10.0	20.0	30.0	40.0	50.0	60.0	70.0	80.0	90.0	100.0

3.3 SENSOR CHARACTERISTICS

Dynamic range

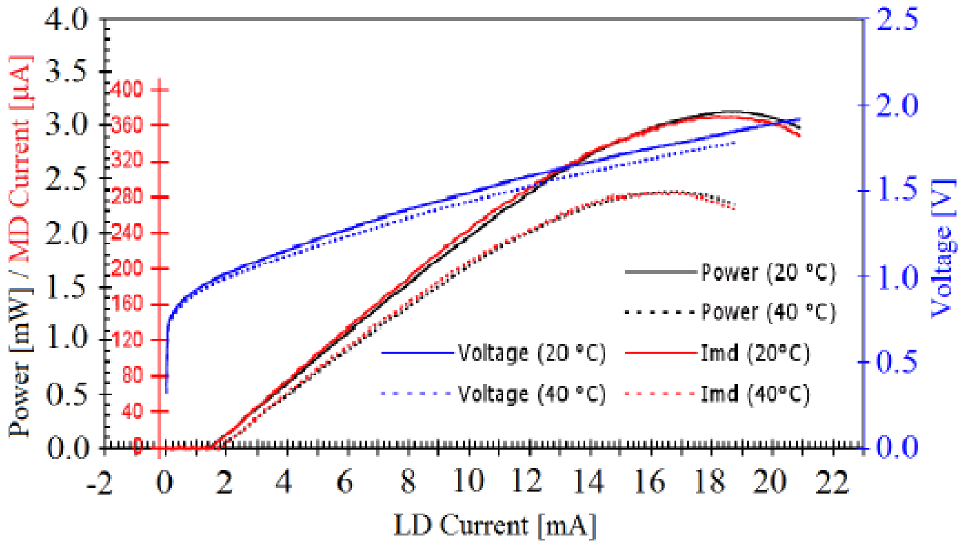
The term dynamic range is also used in more general acceptances.

In the case of a gas sensor, the dynamic range in concentrations represents the entire concentration range within which the sensor has a linear response (transfer function).

In the case of a semiconductor laser source, the current dynamic range represents the entire range of current values within which the laser emits light.

The **Full-Scale Output** is often also specified together with the dynamic range.

It is strictly defined as the difference between the **electrical signals measured when the maximum and minimum stimulus are applied** as input.



3.3 SENSOR CHARACTERISTICS

Accuracy

A very important characteristic of a sensor is **accuracy, which really means inaccuracy**.

Inaccuracy is measured as a **highest deviation of a value represented by the sensor from the ideal or true value** of a stimulus at its input. The true value is attributed to the input stimulus and accepted as having a specified uncertainty because one never can be absolutely sure what the true value is.

The deviation from an ideal (true) transfer function can be described as a **difference between the value that is computed back from the output, and the actual input stimulus value.** That is why inaccuracy is always defined with respect to the dynamic range.

For example, if we repeat this experiment over and over again without any random error and every time we observe an error of 0.5 mm, we may say that the sensor has a systematic inaccuracy of 0.5 mm over a 10-mm span.

3.3 SENSOR CHARACTERISTICS

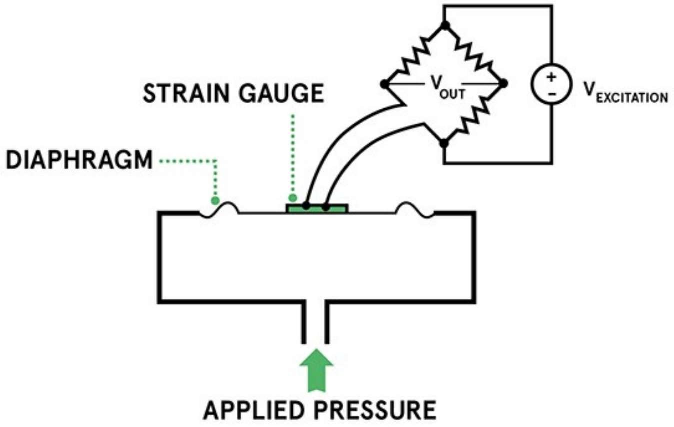
Accuracy

The inaccuracy rating may be represented in a number of forms:

- Directly in terms of measured value (Δ)
- In percent of input span (full scale)
- In terms of output signal

For example, a piezoresistive pressure sensor has a 100-kPa input full scale and a 10 Ω full-scale output. Its inaccuracy may be specified as:

- $\Delta = \pm 500\ Pa$
- In percent of input span : ±0.5%
- In terms of output signal : $\pm 0.05\ \Omega$



3.3 SENSOR CHARACTERISTICS

Accuracy vs. Precision

Precision and accuracy represent two different evaluation of measurement errors.

Accuracy refers to **how close** a measurement is **to the true or accepted value**.

Precision refers to **how close** measurements of the **same item are to each other**.

The two parameters are typically independent from each other. That means it is possible to be very precise but not very accurate, and it is also possible to be accurate without being precise.

Accurate
Precise

Not Accurate
Precise

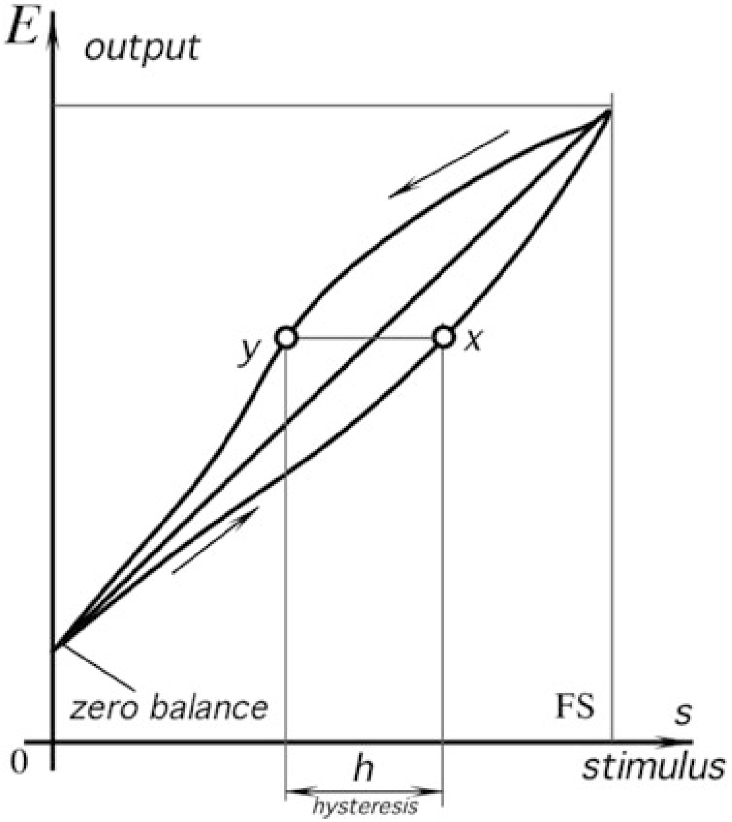
Accurate
Not Precise

Not Accurate
Not Precise

3.3 SENSOR CHARACTERISTICS

Hysteresis

A hysteresis error is a **deviation** of the sensor's output at a specified point of the input signal when it is **approached from the opposite directions**.



For example, a displacement sensor when the object moves from **left to right** at a certain point produces voltage, which differs by 20 mV from that when the object **moves from right to left**.

If the sensitivity of the sensor is 10 mV/mm, the hysteresis error in terms of displacement units is 2 mm.

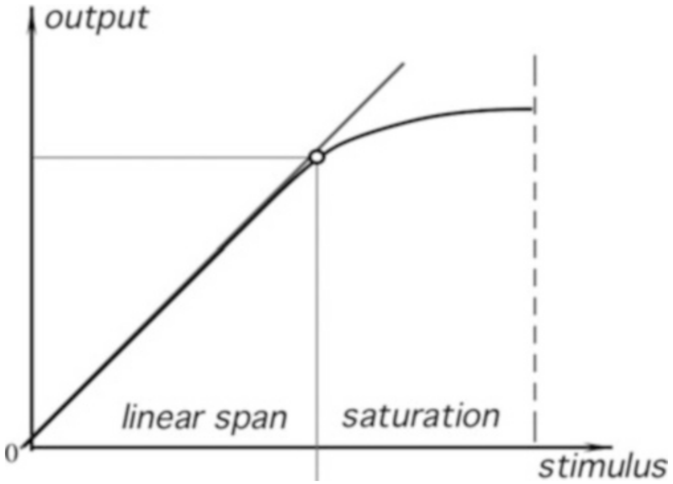
Typical causes for hysteresis are **friction and structural changes in the materials**.

3.3 SENSOR CHARACTERISTICS

Saturation

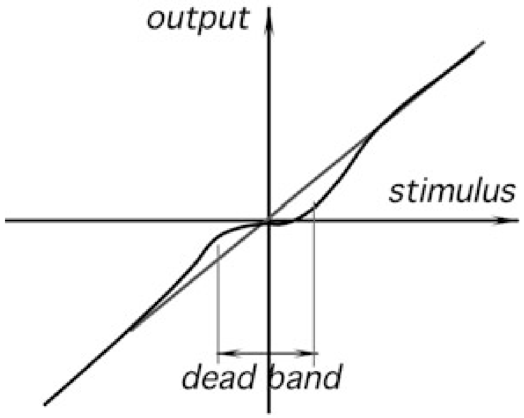
Every sensor has its operating limits. Even if it is considered linear, at some levels of the input stimuli, its output signal no longer will be responsive.

Further increase in stimulus does not produce a desirable output. It is said that the sensor exhibits a span-end nonlinearity or saturation. In these cases, the sensor saturation is reached.



The **dead band is the insensitivity of a sensor in a specific range of input signals**.

In that range, the output may remain near a certain value (often zero) over an entire dead-band zone.



3.3 SENSOR CHARACTERISTICS

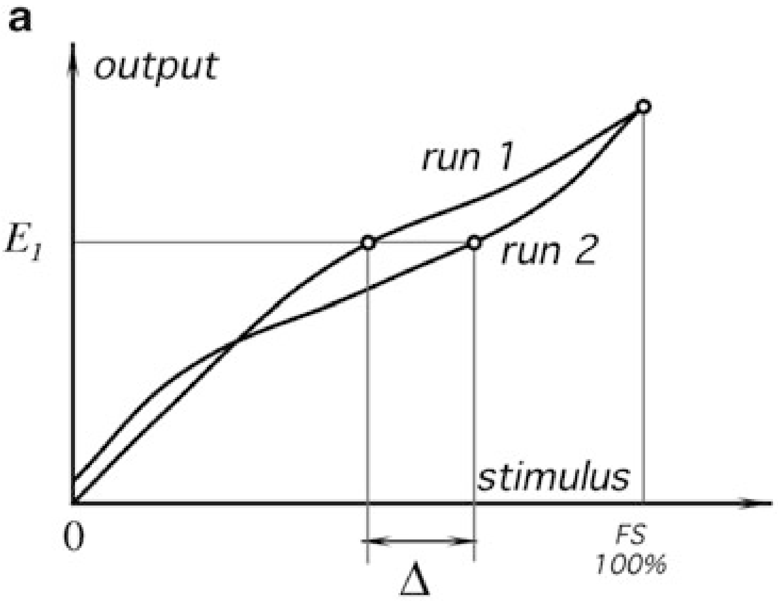
Repeatability

Repeatability (reproducibility) error δ_R is caused by the inability of a **sensor to represent the same value under presumably identical conditions**.

The repeatability is expressed as a maximum difference between the output readings as determined by two run cycles, unless otherwise specified. It is usually represented as percentage of FS:

$$\delta_R = \frac{\Delta}{FS} 100\% \tag{3.12}$$

The possible sources of a repeatability error may be thermal noise, build up charge, material plasticity, etc.



3.3 SENSOR CHARACTERISTICS

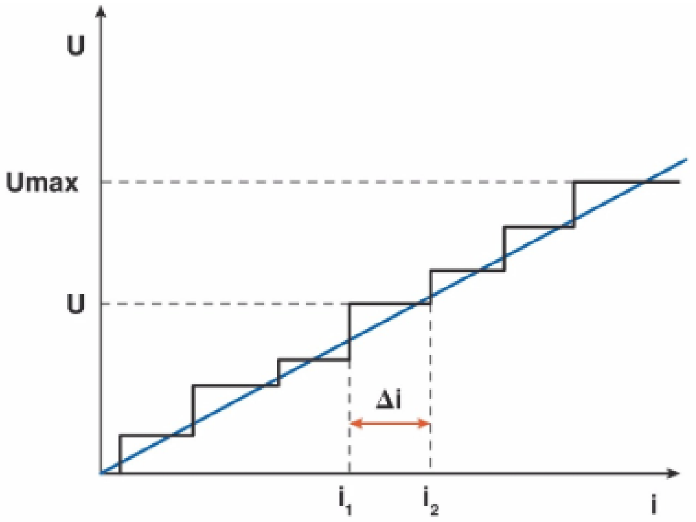
Resolution

Resolution describes the **smallest increment of a stimulus that can be sensed**. When a stimulus continuously varies over the range, the output signals of some sensors will not be perfectly smooth, even under the no-noise conditions. The output may change in small steps.

This parameter is crucial when the signal is converted into a digital format, as it is broken into small steps where a number is assigned to each step. The magnitude of the input variation, which results in the output smallest step, is specified as a resolution under specified conditions.

Sometimes, **it may be specified as percent of a full scale (FS).** For instance, for an angular sensor having 270° FS, a 0.5° resolution may be specified as 0.18 % of FS.

It should be noted that the **step size may vary over the range**, hence, the resolution may be specified as typical, average, or worst.



3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

Under the **static conditions** (a very slow-changing input stimulus) a **sensor is described by a time-invariant** transfer function, accuracy, span, calibration, etc.

However, when an input stimulus varies with an appreciable rate, a sensor response generally does not follow it with a perfect fidelity. The reason is that **both the sensor and its coupling with the stimulus source cannot always respond instantly**.

In other words, a sensor may be described by a time-dependent characteristic that is called **dynamic characteristic**.

If a sensor does not respond instantly, it may represent the stimulus as somewhat different from the real, that is, **the sensor responds with a dynamic error**. A difference between a static and dynamic error is that **the latter is always time dependent**.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

The dynamic characteristic is related to the warm-up time characteristic of the sensor, even if they represent two different specifications.

Warm-up time is a time delay between applying power to the sensor or the excitation signal and the moment when the sensor can operate within its specified accuracy.

Many sensors have a negligibly short warm-up time. However, some sensors, especially those that operate in a thermally controlled environment (a thermostat, e.g.) and many chemical sensors employing heaters, may require seconds and even minutes of a warm-up time before they are fully operational within the specified accuracy limits.

Conversely, **dynamic characteristics represent the response of the sensor to an instantaneous excitation during the whole operating time**.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

In the control theory, it is common to describe the input-output relationship through a **constant-coefficient linear differential equation**. Then, the sensor's dynamic (time-dependent) characteristics can be studied by evaluating such an equation. Depending on the sensor design, differential equations can be of several orders.

A **zero-order sensor** is characterized by the relationship, for a linear transfer function, where the input $s(t)$ and output $E(t)$ are functions of time t:

$$E(t) = a + bs(t) \tag{3.13}$$

where a is called an offset and b is called static sensitivity.

A zero-order sensor requires that the **sensor does not incorporate any energy storage device**, like a capacitor or mass. **It responds instantaneously. In other words, such a sensor does not need any dynamic characteristics**.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

A **First-Order differential equation describes a sensor that incorporates one energy storage component**. The relationship between the input s(t) and output E(t) is a first-order differential equation:

$$b_1 \frac{dE(t)}{dt} + b_o\, E(t) = s(t) \tag{3.14}$$

A typical example of a first-order sensor is a temperature sensor for which the energy storage is thermal capacity.

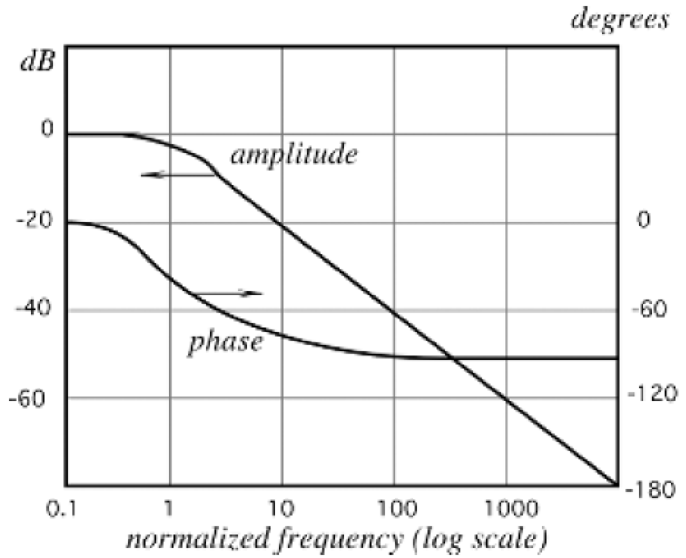
Typically, a dynamic characteristic of a first-order sensor is expressed as a **frequency response**, which specifies how fast the sensor can react to a change in the input stimulus. The frequency response is expressed in hertz or rads per second to specify the relative reduction in the output signal at a certain frequency, thus expressed in decibel unit.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

A commonly used reduction number (frequency limit) is -3 dB. It shows at what frequency the output voltage (or current) drops by about 30 %.

The frequency response limit f_u is often called the **upper cutoff frequency**, as it **is considered the highest frequency that a sensor can process.**



3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

The frequency response directly relates to a **speed response**, which is defined in units of input stimulus per unit of time. Another way to specify speed response is **by time, which is required by the sensor to reach 90% of a steady-state or maximum level upon exposure to a step stimulus**.

For the first-order sensors, the response $E(t)$ as a function of time upon a step-like stimulus is:

$$E(t) = E_m\left(1 - e^{-\frac{t}{\tau}}\right) \tag{3.15}$$

where $\boldsymbol{E_m}$ **is steady-state output** and $\boldsymbol{\tau}$ **is the time constant**. As defined, τ is a measure of the sensor's inertia.

In electrical terms, it is equal to the product of electrical capacitance C and resistance R, $\boldsymbol{\tau = RC}$. In thermal terms, thermal capacity and thermal resistances should be used instead.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

When $\boldsymbol{t} = \boldsymbol{\tau}$, results:

$$\frac{S}{S_m} = 1 - \frac{1}{e} = 0.632 \tag{3.16}$$

Meaning that **after an elapse of time equal to one time constant, the response reaches about 63% of its steady-state level**. Similarly, it can be shown that after two-time constants, the height will be 86.5% and after three-time constants it will be 95%.

In the same way, a **lower cutoff frequency** can be defined as the one indicating the lowest frequency of stimulus that the sensor can process.

The upper cutoff frequency shows how fast the sensor reacts, while the lower cutoff frequency shows how slow the sensor can process changing stimuli

3.3 SENSOR CHARACTERISTICS

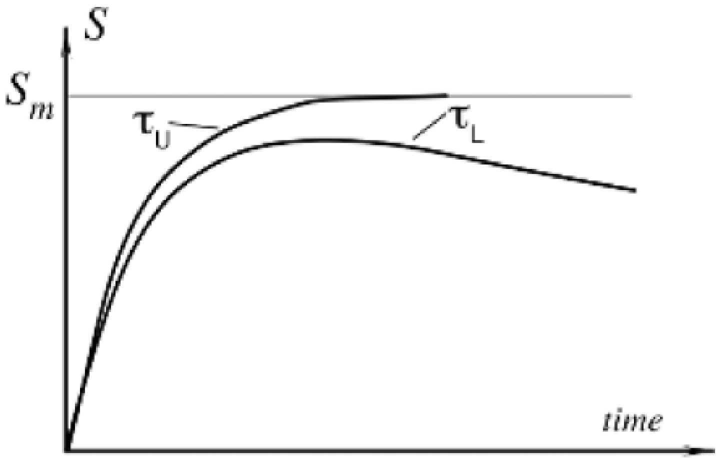
Dynamic characteristics

A time constant can also be defined for the **lower cutoff frequency**. With a higher and a lower cut-off frequency, for a first-order sensor, its response to a step stimulus will be:

$$E(t) = E_m\left(1 - e^{-\frac{t}{\tau_u}}\right)e^{-\frac{t}{\tau_L}} \tag{3.17}$$

where τ_u is the time constant of the upper cutoff frequency while τ_L is the time constant of the lower cutoff frequency.

The sensor never reaches the steady-state response. Sensors in which the two cut-off frequencies are very similar to each other are narrow band sensors. In this case, it is impossible to separate the two cutoff frequencies. For broadband sensors, both time constants can be measured accurately ($\tau_L >> \tau_u$).



3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

A **Second-Order differential equation describes a sensor that incorporates two energy storage components**. The relationship between the input s(t) and output E(t) is a differential equation:

$$b_2 \frac{d^2E(t)}{dt^2} + b_1 \frac{dE(t)}{dt} + b_o\, E(t) = s(t) \qquad (3.18)$$

This is the differential equation describing a forced and damped harmonic oscillator.

A typical example of a second-order sensor is an accelerometer that incorporates an inertial mass and a spring.

A **second-order response** is specific for a sensor that responds with a **periodic signal**.

3.3 SENSOR CHARACTERISTICS

Dynamic characteristics

The output E(t) as a function of time for a step stimulus is:

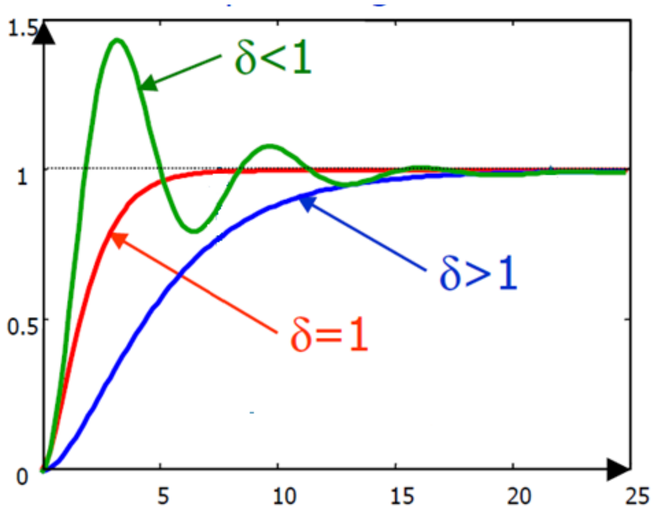
$$E(t) = 1 - Ae^{-\delta\omega_n t}\, sen(\omega t + \varphi) \qquad (3.19)$$

where δ **is the damping coefficient** and ω_n **the natural angular frequency**:

$$A = \frac{1}{\sqrt{1-\delta^2}};\ \omega = \omega_n\sqrt{1-\delta^2};\ \varphi = arctg\left(\frac{\sqrt{1-\delta^2}}{\delta}\right) = arcsen\sqrt{1-\delta^2}; \qquad (3.20)$$

The value of the damping coefficient determines the sensor behavior:

- If $\delta \geq 1$, the sensor response is comparable to that of a first order system, and the system is overdamped or critically damped.
- if $\delta < 1$, the sensor response shows oscillations with decreasing amplitude over time before reaching the steady-state value.



3.4 PHYSICAL PRINCIPLES OF SENSING

Physics underlying sensing

A sensor is a converter of generally nonelectrical effects into electrical signals, thus one and often several transformation steps are required before the electric output signal can be generated.

These steps involve changes of types of energy or physical properties of materials, wherein the final step shall produce electrical signal of a desirable format.

There are several physical effects that result in a direct generation of electrical signals in response to nonelectrical influences and thus can be used in sensors.

- **Capacitance effect**
- **Resistance effect**
- **Piezoelectric effect**
- **Pyroelectric effect**
- **Thermoelectric effects**

3.4 PHYSICAL PRINCIPLES OF SENSING

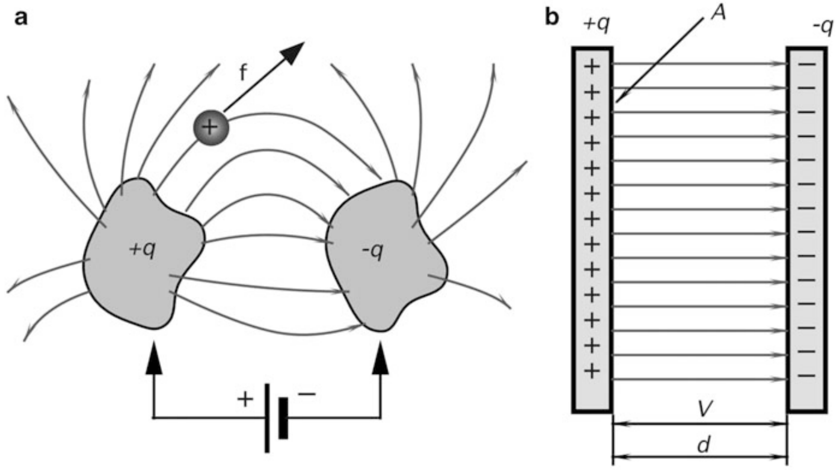
Capacitance effects

Two isolated conductive objects of arbitrary shape connected to the opposite poles of a battery will receive equal amounts of opposite charges.

That is, a negatively charged plate will receive additional electrons while there will be a deficiency of electrons in the positively charged plate.

Now, let us disconnect the battery. If the plates are totally isolated and exist in a vacuum, they will remain charged theoretically infinitely long. **A combination of the plates, which can hold an electric charge, is called a capacitor**.

As a result, an electric field is generated among the plates of the capacitor, exiting the positive charged plate and pointing towards the negative charged one.



3.4 PHYSICAL PRINCIPLES OF SENSING

Capacitance effects

The capacitor may be characterized by *q*, the magnitude of charge on either conductor, and by *V*, the positive potential difference between the conductors.

The charge *q* is not a net charge on the capacitor, which is zero.

V is not the potential of either plate, but the potential difference between them.

The ratio of charge to voltage is constant for each capacitor:

$$\frac{q}{V} = C \tag{3.21}$$

This fixed ratio, **C is called the capacitance of the capacitor**. Its value depends on the shapes and relative position of the plates.

The ratio C also depends on the medium in which the plates are immersed.

3.4 PHYSICAL PRINCIPLES OF SENSING

Capacitance effects

This capacitance effect is widely employed in sensing.

In a capacitive sensor, capacitance is modulated (modified) by an external stimulus or by a signal from an intermediate transducer. Thus, **to vary capacitance, the stimulus needs to change one of the parameters that define the capacitance**.

For example, for a flat-plate capacitor placed in vacuum results:

$$C = \frac{\varepsilon_0 A}{d} = \varepsilon_0 G \tag{3.22}$$

Where ε_0 is the dielectric constant of vacuum, *A* is the area of the flat plate and *d* is their mutual distance. G is called geometric factor of the capacitor.

Eq. (3.22) show the relationship between the plate area and distance between the plates. Varying one of them will change the capacitance affecting in turn the voltage V. **This equation holds for a capacitor in vacuum or air.**

3.4 PHYSICAL PRINCIPLES OF SENSING

Capacitance effects

When a dielectric medium is placed between the capacitor’s plates it increase the capacitance of the device by a factor of ***κ***, which is known as the **dielectric constant of the material**.

The increase in capacitance due to the dielectric presence is a result of molecular polarization, reducing the electric field in the capacitor, and thus reducing its voltage, of a factor κ.

$$C = \kappa \frac{q}{V_0} = \kappa C_0 = \varepsilon_0 \kappa G \tag{3.23}$$

The dielectric constant depends on material, temperature, and humidity/moisture content. **All these variables can be used in the capacitive sensors as inputs to modulate capacitance**.

An example of a capacitive sensor is a humidity sensor. In such a sensor, a dielectric filling between the capacitor plates is fabricated of a material that is hygroscopic, that is, it can absorb water molecules. The material dielectric constant varies with the amount of absorbed moisture.

3.4 PHYSICAL PRINCIPLES OF SENSING

Resistance effects

A material is characterized by its ability to pass electric current, called **resistivity**. The material is said to be an Ohmic resistor if its electrical resistance is defined by Ohm's law, meaning that a ratio of voltage to current is a constant.

$$R = \frac{V}{i} \tag{3.24}$$

Resistance is a characteristic of a device. It depends on both: the material type and geometry of the resistor. Material itself can be characterized by a **specific resistivity, ρ**.

However, the **specific resistivity of a material is not constant**. It **changes with temperature**, and in a relatively narrow temperature range may be linearly approximated through the thermal sensitivity (slope) α, which is the temperature coefficient of resistivity:

$$\rho = \rho_0 \left(1 + \alpha \frac{T - T_0}{T_0}\right) \tag{3.25}$$

where ρ_0 is the specific resistivity at reference temperature T_0 (commonly T_0=0°C or 25°C).

3.4 PHYSICAL PRINCIPLES OF SENSING

Resistance effects

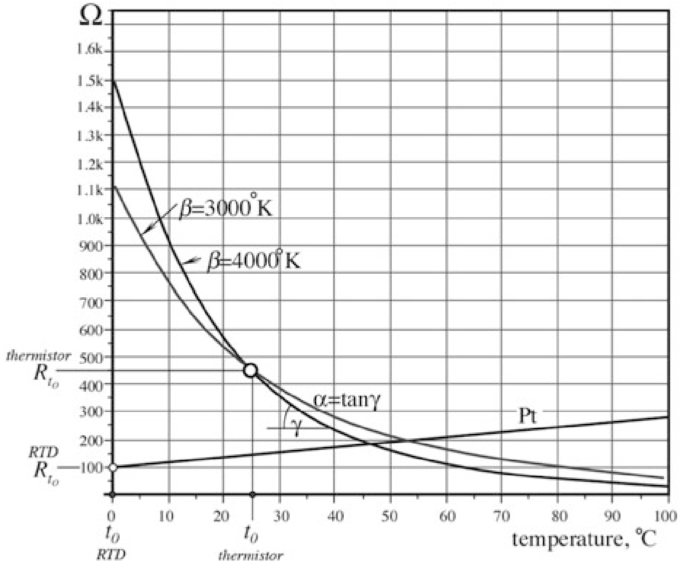
A strong α allows fabricating two types of temperature sensors: **resistance temperature detector (RTD) and thermistors.**

The most popular RTD is a platinum (Pt) sensor that operates over a broad temperature range up to 600 °C. It’s transfer function can be approximated as:

$$R = R_0(1 + 39.08 \times 10^{-4}T - 5.8 \times 10^{-7}T^2)\Omega \tag{3.26}$$

Thermistors are ceramic semiconductors commonly made of oxides of one or more of the following metals: nickel, manganese, cobalt, titanium, iron. Oxides of other metals are occasionally used. Thermistors possess nonlinear temperature-resistance characteristics, and the most popular of the thermistor’s transfer function approximations is the exponential form:

$$R = R_0 e^{\beta\left(\frac{1}{T} - \frac{1}{T_0}\right)} \tag{3.27}$$



3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

The piezoelectric effect is the generation of electric charge by a crystalline material upon subjecting it to stress.

The effect exists in natural crystals, such as quartz (SiO_2), and poled (artificially polarized) man-made ceramics and some polymers, such as polyvinylidene flouride.

The piezoelectric effect was discovered Curie brothers (Pierre Curie e Paul-Jacques Curie) discovered the piezoelectric effect in quartz in 1880.

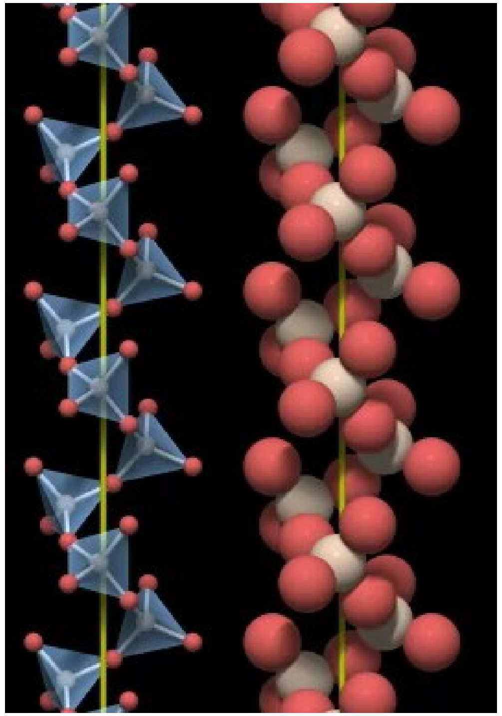
However, very little practical use was made until 1917, when another Frenchman, Professor P. Langevin used x-cut plates of quartz to generate and detect sound waves in water. His work led to the development of sonar.

3.4 PHYSICAL PRINCIPLES OF SENSING

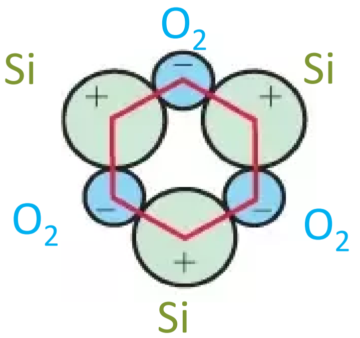
Piezoelectric effect

A simplified, yet quite explanatory model of the piezoelectric effect was proposed in 1927 by A. Meissner.

A quartz crystal is modeled as a helix with one silicon (Si) and two oxygen (O_2) atoms alternating around the helix.



A quartz crystal is cut along its axes x, y, and z; in Figure there is a view along the z-axis. In a single-crystal cell, there are three silicon atoms and six oxygen atoms. Oxygen is being lumped in pairs.



Each silicon atom carries four positive charges, and a pair of oxygen atoms carries four negative charges (two per atom). **Therefore, a quartz cell is electrically neutral under the no-stress conditions.**

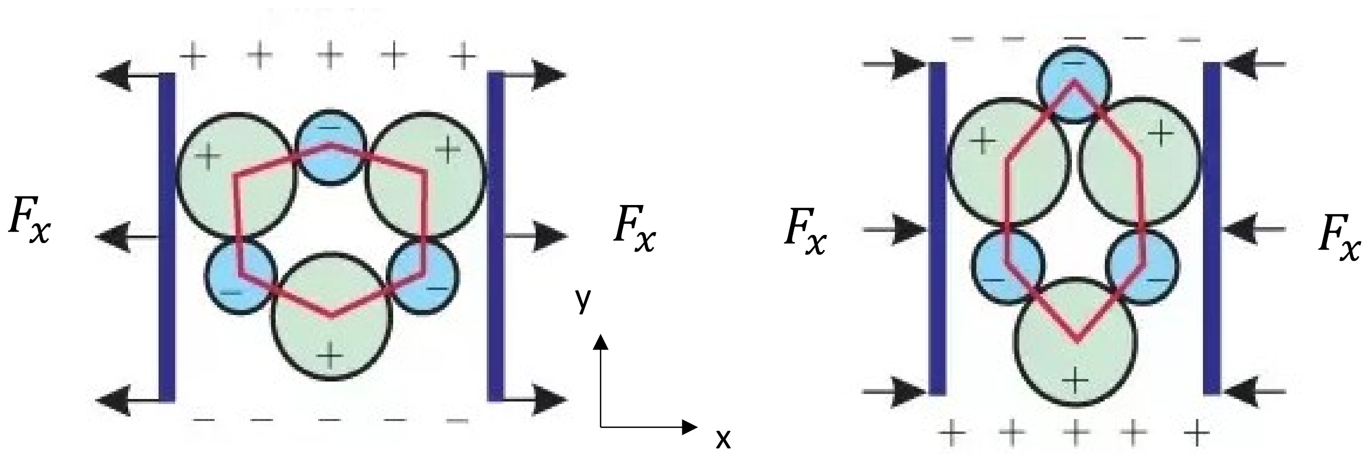
3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

When an external force (F_x) is applied along the x -axis, the hexagonal lattice becomes deformed.

A **compressing force** which shifts atoms in a crystal in such a manner that a positive charge is built up at the silicon atom side and a negative charge at the oxygen pair side. Thus, the crystal develops an electric charge along the y-axis.

If the **crystal is stretched along the x-axis**, a charge of opposite polarity is built along the y-axis, which is a result of a different deformation.



3.4 PHYSICAL PRINCIPLES OF SENSING

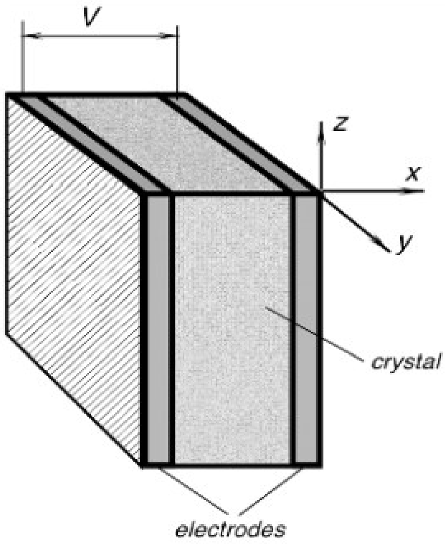
Piezoelectric effect

This simple model illustrates that crystalline material can develop **electric charge on its surface in response to a mechanical deformation**.

To pick up an electric charge, conductive electrodes must be applied to the crystal at the opposite sides of the cut. As a result, a **piezoelectric sensor becomes a capacitor with a dielectric material which is a piezoelectric crystal**. The dielectric acts as a generator of electric charge, resulting in voltage V across the capacitor.

Although charge in a crystalline dielectric is formed at the location of an acting force, **metal electrodes equalize charges along the surface, making the capacitor not selectively sensitive**.

However, if **electrodes are formed with a complex pattern**, it is possible to determine the exact location of the applied force by measuring the response from a selected electrode.



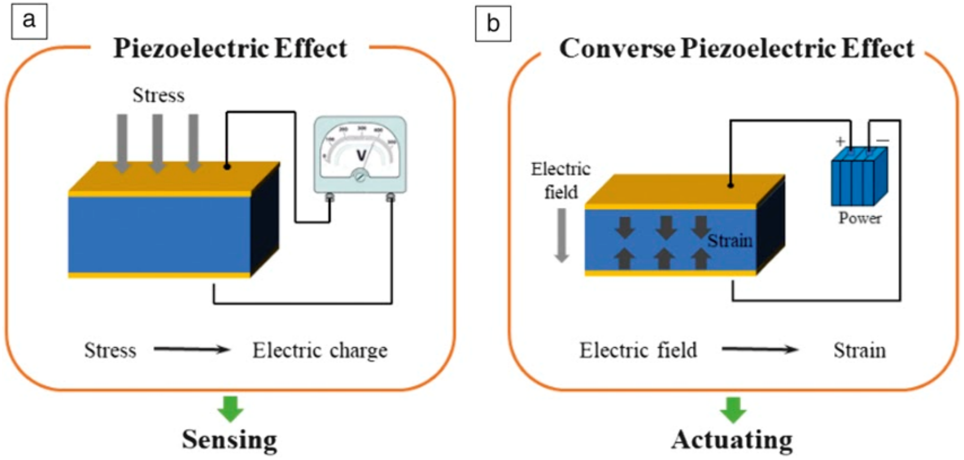
3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

The **piezoelectric effect is a reversible physical phenomenon**. That means that applying voltage across the crystal produces mechanical strain. Therefore, we can define:

- **Direct piezoelectric effect**: crystal deformation induces charges generation
- **Inverse piezoelectric effect**: external applied voltage induces crystal deformation

By placing several electrodes on the crystal, it is possible to use one pair of electrodes to deliver voltage to the crystal and the other pair of electrodes to pick up charge resulting from developed strain. This method is used quite extensively in various piezoelectric transducers.



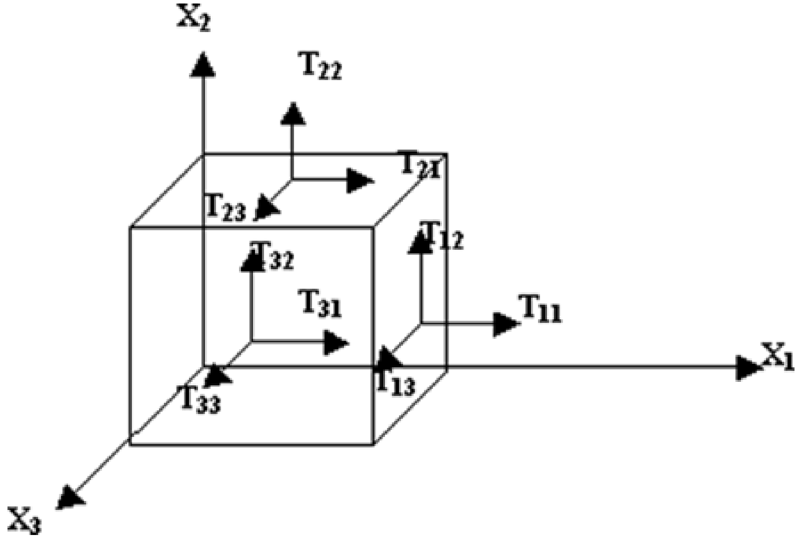
3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

The description of the piezoelectric phenomenon and its properties requires the preliminary definition of some mechanical quantities involved: **the mechanical stress and the strain**.

Mechanical stress is the uniform pressure applied on a crystal along one or more directions. It is expressed in [N/m^2] and as a tensor T_{ij} depending on the direction and surface on which the stress is applied.

In crystallography, the axes are indicated as x_1, x_2 and x_3. The first subscript of the tensor T_{ij} identifies the plane (through its orthogonal axis) while the second subscript the direction.



3.4 PHYSICAL PRINCIPLES OF SENSING

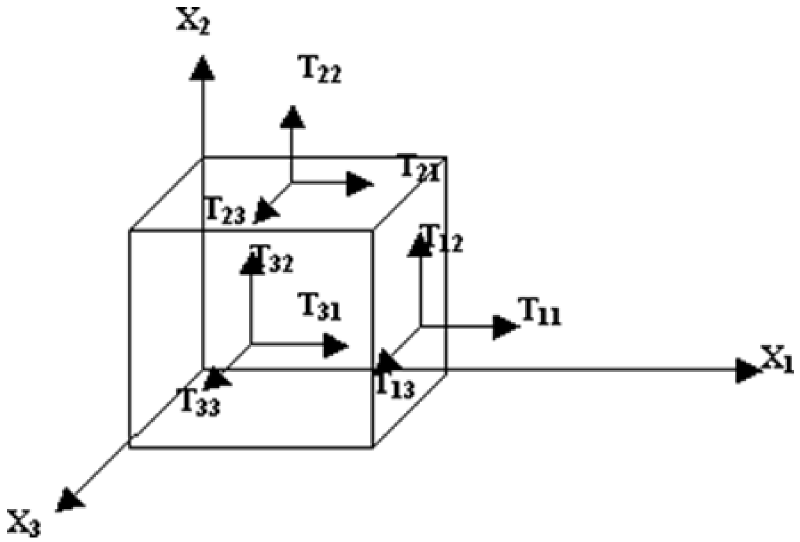
Piezoelectric effect

A **piezoelectric crystal is a rigid body**, thus the stress components can be related to each other.

$$T_{12} = T_{21}$$

$$T_{23} = T_{32}$$

$$T_{13} = T_{31} \qquad (3.28)$$



For this reason, this convention is often used in solids to simplify the double-subscript notation:

$$\begin{array}{ll} T_1 = T_{11} & T_4 = T_{23} \\ T_2 = T_{22} & T_5 = T_{13} \\ T_3 = T_{33} & T_6 = T_{12} \end{array} \qquad (3.29)$$

3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

Strain is the deformation of a material from stress. It is simply a ratio of the change in length to the original length. It is a dimensionless and represented by the tensor S_i with i=1, 2, …, 6, similarly to the case of stress tensor.

There is a **direct proportionality** between T_i and S_i expressed by the relation:

$$T_i = Y_{ij}^{E} S_j \tag{3.30}$$

where $Y_{ij}{}^{E}$ **is the constant of proportionality, better known as Young's modulus**. This constant depends on the characteristics of the material and in particular on the density and propagation speed of an acoustic wave within the medium. The **inverse relation** is:

$$S_j = s_{ij} T_i \tag{3.31}$$

which allows the definition of the other constant of proportionality s_{ij}**, known as compliance**.

3.4 PHYSICAL PRINCIPLES OF SENSING

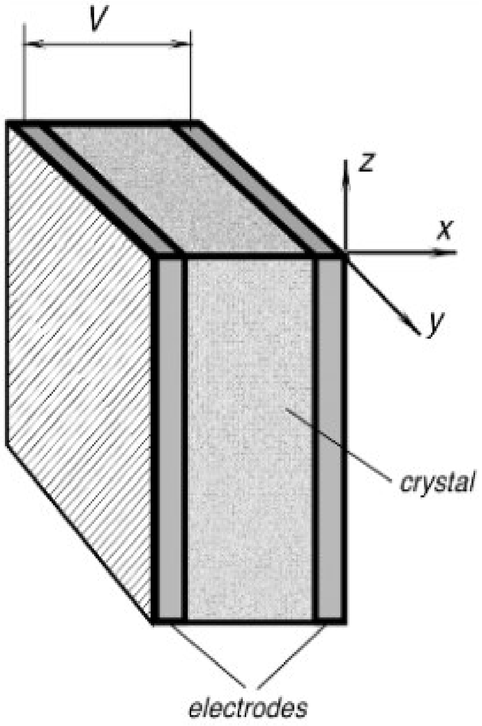
Piezoelectric effect

If a stress field is applied on a crystal, a strain field is generated, resulting in an **accumulation of surface charges** between two, **which in turns produces an electric field $\vec{\mathbf{E}}$.**

Being σ_l the surface charge density generated on a plate in the absence of dielectric and $-\sigma_l$ that generated on the other plate, the electric field is:

$$\overrightarrow{E_0} = \frac{\sigma_l}{\varepsilon_0}\widehat{u_0} \tag{3.32}$$

with $\widehat{u_0}$ which identifies direction of $\overrightarrow{E_0}$ and ε_0 is the vacuum permittivity.



3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

Since the dielectric is now polarized, the effective net charge density is reduced and the electric field $\overrightarrow{\Delta E}$ due to electrical polarization, being $\vec{P}$ the polarization vector:

$$\overrightarrow{\Delta E} = -\frac{\vec{P}}{\varepsilon_0} \qquad (3.33)$$

Considering the sum of both the contributions:

$$\vec{E} = \overrightarrow{E_0} + \overrightarrow{\Delta E} \qquad (3.34)$$

Replacing Eq. (3.33) and (3.34) in (3.32), results:

$$\sigma_l = \varepsilon_0 E + P \qquad (3.35)$$

where the moduli of $\vec{E}$ and $\vec{P}$ are considered, because the vectors are supposed to be parallel.

Thus, when the electric potential is fixed, the free charge on a conductor surrounded by a dielectric is a function of the polarization of the dielectric itself.

3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

You can now define the displacement vector field $\vec{D}$ through its three Cartesian components:

$$D_i = \varepsilon_0 E_i + P_i \tag{3.36}$$

In the case of piezoelectric material, the polarization causing the electrical displacement it is due to the applied mechanical stress.

The **piezoelectric charge constant, $\boldsymbol{d_{ij}}$** (with $i = 1, 2, 3$ and j $= 1, 2, \dots, 6$), is defined as the electrical polarization induced in a material per unit of mechanical stress applied.

Consequently, in the **direct piezoelectric effect**, the electrical displacement is directly proportional to the applied stress:

$$D_i = d_{ij}\, T_j \tag{3.37}$$

In the case of **inverse piezoelectric effect**, the proportionality is still valid and the relation between the electric field and the strain field is:

$$S_j = d_{ji}\, E_i \tag{3.38}$$

3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

The constant of proportionality is identical to that introduced in the direct effect and it is denoted again with d_{ji}, even if in this case it denotes the mechanical strain on the material per unit of electric field applied.

In matrix form, $D_i = d_{ij}\, T_j$ becomes:

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} \tag{3.39}$$

At room temperature, for a **z-cut quartz crystal**, the piezoelectric matrix is:

$$\begin{bmatrix} d_{11} & -d_{11} & 0 & d_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -d_{14} & -2d_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.40}$$

Where $\boldsymbol{d_{11} = 2.3 \cdot 10^{-12}\ C/N}$ and $\boldsymbol{d_{14} = 0.67 \cdot 10^{-12}\ C/N}$, are the values reported in literature.

3.4 PHYSICAL PRINCIPLES OF SENSING

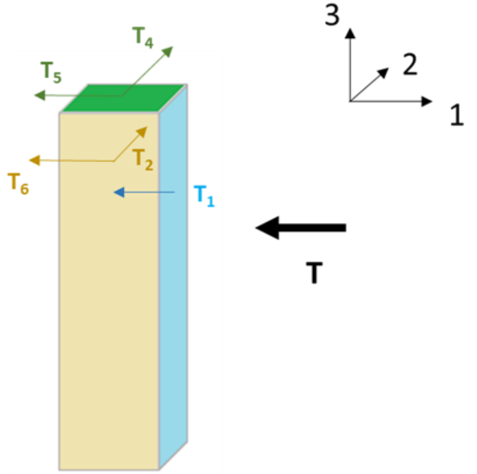
Piezoelectric effect

Thus, the components of the electric displacement vector for a quartz plate is:

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} d_{11} & -d_{11} & 0 & d_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -d_{14} & -2d_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{bmatrix} \Rightarrow \begin{aligned} D_2 &= -d_{14}T_5 - 2d_{11}T_6 \\ D_1 &= d_{11}T_1 - d_{11}T_2 + d_{14}T_4 \end{aligned} \qquad (3.41)$$

Consider a plate of z-cut quartz, and mechanical stress $\vec{T}$ **applied along the direction 1**:

We indicate the components of mechanical stress that contribute to polarization, using the two expressions found for D_1 and D_2.



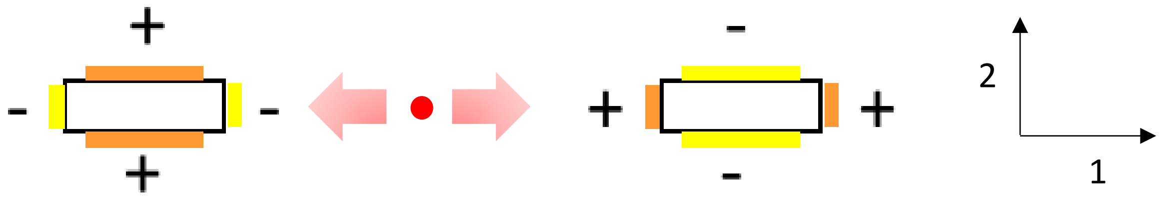
3.4 PHYSICAL PRINCIPLES OF SENSING

Piezoelectric effect

If we assume that the deformations occur only along the direction 1, T_2 and T_4 can be neglected, resulting in:

$$\begin{gathered} D_1 = d_{11} T_1 \\ D_2 = -d_{14} T_5 - 2 d_{11} T_6 \end{gathered} \tag{3.42}$$

This means that polarization charges arise along the direction 1, while other ones of opposite sign along the direction 2. **Therefore, the quartz plate acts as an electric quadrupole.**



For this reason, the in a piezoelectric sensor the electrodes must be appropriately designed in order to collect charges **with the same polarity**.

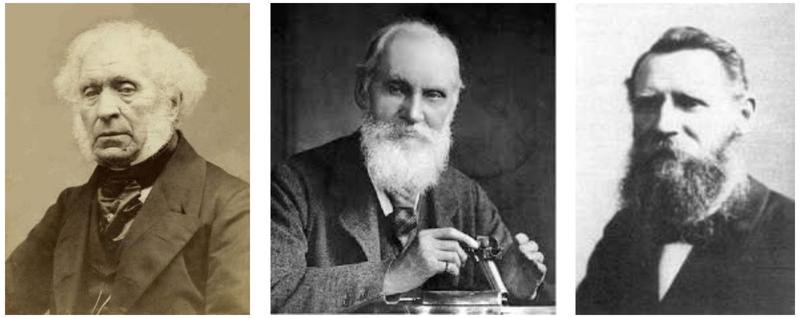
3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

Pyroelectric materials are crystalline substances capable of generating an electrical charge in response to heat flow. A crystal is considered to be pyroelectric if it exhibits a spontaneous temperature dependent polarization.

In 1717 Louis Lemery was the first to describe this phenomenon without giving it the name used today. In 1747 Linneaus relates this phenomenon to electricity. In the 1756, this assertion was supported by scientific observations and therefore by scientific evidence conducted by Franz Ulrich Theodor Aepinus.

In 1824 Sir David Brewster gave the effect the name it has today. Both William Thomson in 1878 and Woldemar Voigt in 1897 helped develop a theory for the processes behind pyroelectricity.



3.4 PHYSICAL PRINCIPLES OF SENSING

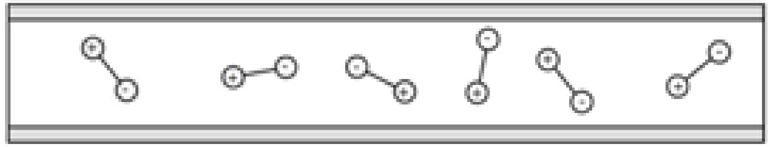
Pyroelectric effect

A pyroelectric material can be considered as a composition of a large number of minute crystallities, each of which behaves as a small electric dipole.

Usually, crystallites can be considered as dipoles that, in some materials, are naturally oriented along the axes of symmetry of the crystal itself as in the case of quartz. In other materials this orientation must be forced by applying a high electric field in such a way as to polarize the material.

There are **different techniques to polarize a material, the most used is based on the thermal polarization** that takes place according to the following phases:

1. The crystal with randomly oriented dipoles is heated to a value below its Curie temperature T_C. The increase of the temperature causes a thermal molecular agitation allowing an easier orientation of the dipoles.



Heating phase, $T < TC$

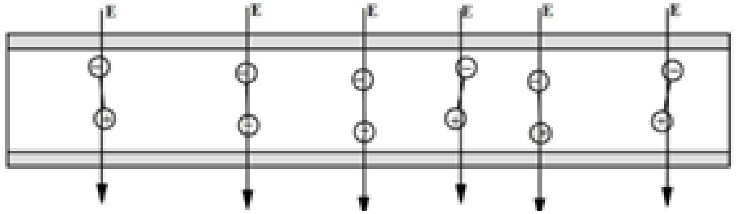
3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

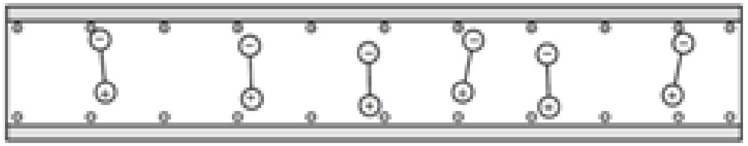
2. An electric field is applied causing dipoles orientation along the field lines. This orientation is usually total, but there may be minimal deviances from the direction of application of the electric field, i.e., the dipoles will be aligned to the electric field with a tolerance range in terms of inclination.

3. The material is cooled while the electric field is still applied.

4. The electric field is removed, and the polarization is considered complete. Until $T < T_C$, the polarization can be considered permanent because the dipoles remain in the position assumed during the process



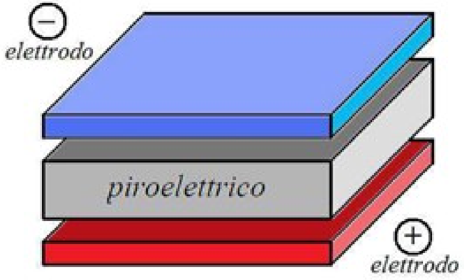
Applied electric field *E*



Pyroelectric material

3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect



By varying the temperature of the pyroelectric material, the polarization varies, and electric charges are generated.

Consider a planar pyroelectric element where the thickness is much less than the other two dimensions. The dipole moment, M, of the bulk pyroelectric sensor is :

$$M = \mu A h \tag{3.43}$$

where μ is the dipole moment per unit volume, A is the sensor's area, and h is the thickness.

The charge, Q_a, which can be picked up by the electrodes, develops the dipole moment across the material:

$$M = Q_a h \tag{3.44}$$

From Eq. (3.43) and (3.44), results: $\boldsymbol{Q_a = \mu A}$

3.4 PHYSICAL PRINCIPLES OF SENSING

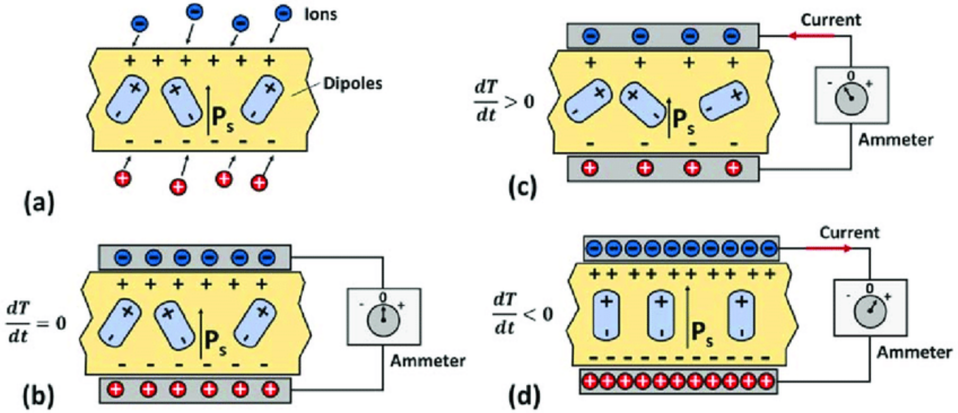
Pyroelectric effect

As the temperature varies, the dipole moment also changes, resulting in an induced charge.

Thermal absorption may be related to a dipole change, so that μ must be considered as a function of both temperature, T_a, and an incremental thermal energy, ΔW, absorbed by the material:

$$\Delta Q_a = A\mu(T_a, \Delta W) \tag{3.45}$$

This means that, **when the pyroelectric material absorbs thermal energy, a charge variation is generated**. This is related to the size of the pyroelectric and the dependence of the dipole moment on the temperature.



3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

Using the same formalism introduced for piezoelectric materials, we can rewrite the same relationships, **including piezoelectric, elastic and thermal phenomena into a single relation**.

The strain field can be expressed as:

$$S_i = s_{ij}T_j + d_{ik}\, E_k + \alpha_i T \tag{3.46}$$

where s_{ji} is the compliance, T_j mechanical stress, d_{jk} piezoelectric coefficients, E_k the electric field, α_i the coefficients of the linear thermal expansion and T the temperature of the crystal, and with $i, k = 1, 2, 3$ and $\mathrm{j} = 1, 2, \dots, 6$).

3.4 PHYSICAL PRINCIPLES OF SENSING

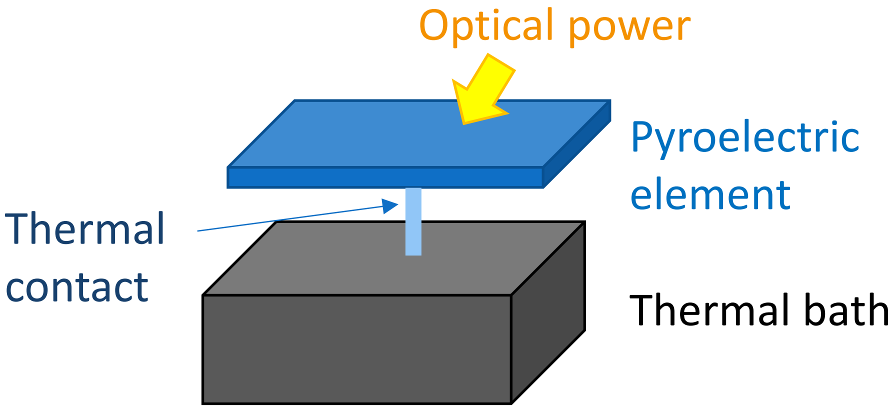
Pyroelectric effect

The main application of pyroelectric materials is as infrared photodetector.

The pyroelectric material is connected by a weak thermal contact with **thermal conductance** $G\ [W/K]$ to a **thermal bath (heat sink) at temperature T_0**. Suppose there are no leak channels for heat dissipation.

If the pyroelectric material absorbs a constant power P_0, the temperature increase of T_1 compared to that of the thermal bath, then the thermal conductance will be:

$$G = \frac{P_0}{T_1} \qquad (3.47)$$



3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

Let’s suppose an increase $P_v\,(t)$ of the optical power due to an optical absorption.

The temperature of the sensing element will change following the time dependence of $P_v(t)$, in relation to its heat capacity C:

$$\eta P_v(t) = \frac{dQ}{dt} = C\,\frac{dT_1}{dt} \tag{3.48}$$

where $\boldsymbol{\eta}$ is the portion of power absorbed by the sensing element (**quantum efficiency**) and the heat capacity defined as $\boldsymbol{dQ} = \boldsymbol{CdT_1}$.

The total power $P_T\,(t)$ absorbed by the sensing element will be:

$$\boldsymbol{P_T(t) = P_0 + \eta P_v(t) = GT_1 + C\,\frac{dT_1}{dt}} \tag{3.49}$$

3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

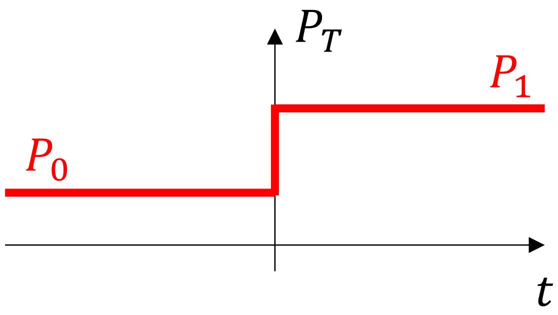
Suppose we **evaluate the response of the sensing element to a step-like stimulus**, thus assuming that:

$$P_T(t) = \begin{cases} P_0 & t < 0 \\ P_1 & t \geq 0 \end{cases} \tag{3.50}$$

P_0 is the power incident on the sensing element (external light), not due to the amount to be measured, P_1.

- For $t < 0$, $P_v(t) = 0$ which means that T_1 is not time-dependent $\left(\frac{dT_1}{dt} = 0\right)$ and the trivial solution is $T_1 = \frac{P_0}{G}$.
- For $t \geq 0$, the differential equation becomes:

$$P_0 + \eta P_1 = GT_1 + C\frac{dT_1}{dt} \tag{3.51}$$



3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

Eq. (3.51) can be solved by imposing the condition of continuity for $t = 0$, $T_1 = \frac{P_0}{G}$. Thus, the complete solution is:

$$T_1(t) = \begin{cases} \frac{P_0}{G} & t < 0 \\ \frac{P_0}{G} + \frac{\eta P_1}{G}\left(1 - e^{-\frac{G}{C}t}\right) & t \geq 0 \end{cases} \tag{3.52}$$

allowing the definition of the thermal time constant for the sensing element:

$$\tau_T = C/G \tag{3.53}$$

For $t \gg \tau_T, T_1 \rightarrow P_0 + \eta P_1$.

Therefore, by measuring $\mathbf{T_1}$ it is possible to calculate the amount of power incident on the sensitive element.

For fast response times, the pyroelectric element must have a low heat capacity and a large thermal conductance with the thermal bath; in other words, it must relax the accumulated excess of heat as quickly as possible.

3.4 PHYSICAL PRINCIPLES OF SENSING

Pyroelectric effect

Now suppose to evaluate the response of the sensing element when exposed to a light beam with an oscillating component at angular frequency ω and amplitude P_ω, together with a DC component P_{DC}:

$$P_v(t) = P_{DC} + P_\omega e^{i\omega t} \tag{3.54}$$

Inserting it into equation (3.51), the amplitude T_ω of the oscillating component of the temperature can be determined:

$$T_\omega = \frac{\eta P_\omega}{\sqrt{G^2+\omega^2 C^2}} \tag{3.55}$$

and the phase shift between the oscillations of the incident radiation field and the temperature oscillations is:

$$\phi = \arctan\left(\frac{\omega C}{G}\right) \tag{3.56}$$

3.4 PHYSICAL PRINCIPLES OF SENSING

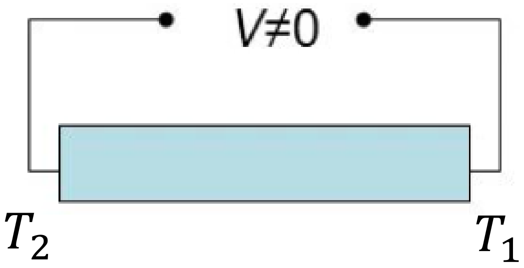
Thermoelectric effect – Seebeck effect

The **Seebeck effect** is a phenomenon in which **a temperature difference between two dissimilar electrical conductors or semiconductors produces a voltage difference** between the two substances.

In 1821, Thomas Johann Seebeck (1770–1831) accidentally joined semicircular pieces of bismuth and copper while studying the thermal effects on galvanic arrangements .

If we take a conductor and place one end of it into a cold place (T_1) and the other end into a warm place (T_2), energy will flow from the warm to cold part. The energy takes the form of heat. The intensity of the heat flow is proportional to the thermal conductivity of the conductor.

In addition, the thermal gradient sets an electric field inside the conductor.



3.4 PHYSICAL PRINCIPLES OF SENSING

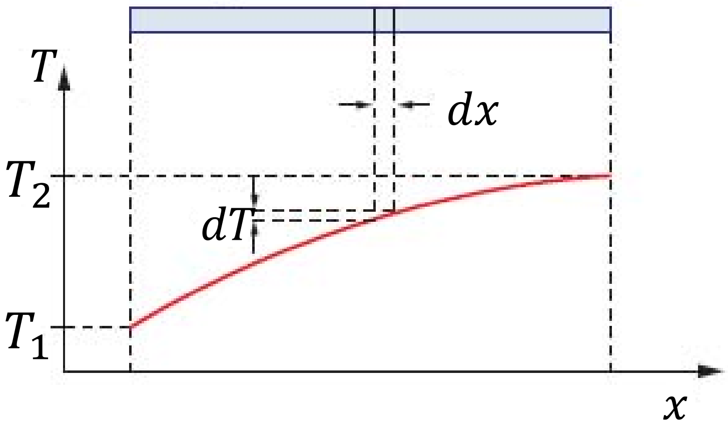
Thermoelectric effect – Seebeck effect

If we denote with dT as the temperature gradient across a small length dx, **being α_a the absolute Seebeck coefficient of the material a**, the potential difference dV_a will be:

$$dV_a = \alpha_a \frac{dT}{dx} dx \tag{3.57}$$

If the material is homogeneous, α_a is not a function of length and the previous equation reduces to:

$$dV_a = \alpha_a dT \tag{3.58}$$



3.4 PHYSICAL PRINCIPLES OF SENSING

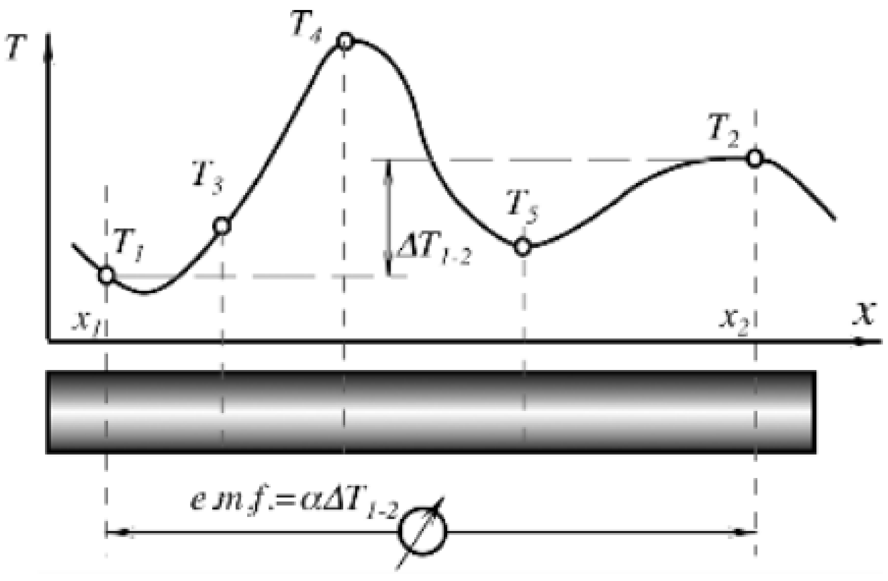
Thermoelectric effect – Seebeck effect

Consider a conductor having **nonuniform temperature T** along its length x.

A **temperature gradient** between any arbitrary points **defines an electromotive force (emf)** between these points.

Then the emf between the two ends of the conductor can be calculated as the sum of the $emfs$ between intermediate points.

Other possible temperatures between the selected points (temperatures T_3, T_4 and T_5, for example) **have no effect** whatsoever on the value of emf between points 1 and 2.



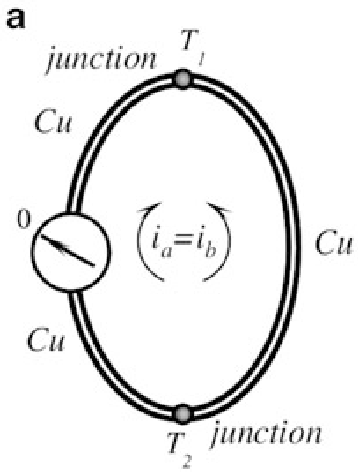
3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Seebeck effect

To measure e.m.f., we would like to connect a voltmeter to the conductor, and this is not as simple as may first look. To measure thermally induced e.m.f. we would need to attach the voltmeter probes. However, the probes are also made of conductors that may be different from the conductor we study. **As a result, the probe contacts will introduce their own e.m.f. and disturb out experiment.**

However, even employing a voltmeter made of the same material of the conductor, the problem is not solved. Let us consider a simple measurement electric circuit where a current loop is formed. We cut the left side of the conductor (Cu) and insert the current meter into the cut in series with the wire.

If the entire loop is made of a uniform material, then no current will be observed even if the temperature along the conductor is not uniform. **Electric fields in the left and right arms of the loop produce equal currents $i_a = i_b$ which cancel one another, resulting in zero net current.** A thermally induced e.m.f. exists in every thermally nonhomogeneous conductor, but it cannot be directly measured.



3.4 PHYSICAL PRINCIPLES OF SENSING

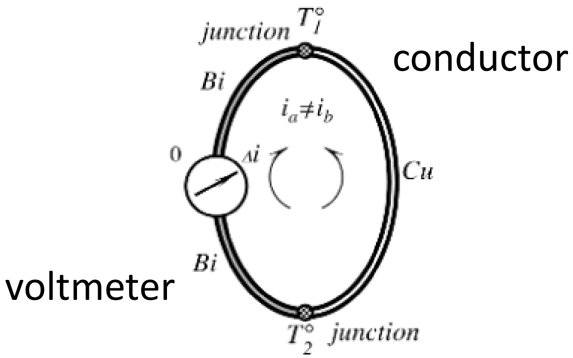
Thermoelectric effect – Seebeck effect

Therefore, in order to observe thermoelectricity, **it is necessary to have a circuit composed of two different materials**, and we can then measure the net difference between their thermoelectric properties.

A loop of two dissimilar metals, placed at two different temperatures produces net current:

$$\Delta i = i_a - i_b \tag{3.59}$$

If, on the other hand, instead of current we measure the net voltage across the broken conductor, the potential will depend only on the materials and the temperature difference.



Using the relation $dV_i = \alpha_i dT$, the **net voltage V_N** will be:

$$V_N = \int_{T_1}^{T_2} \alpha_a dT + \int_{T_2}^{T_1} \alpha_b dT = \int_{T_1}^{T_2} (\alpha_a - \alpha_b) dT \tag{3.60}$$

3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Seebeck effect

When a combination of two dissimilar materials (a and b) is used, the Seebeck potential is determined from a differential Seebeck coefficient:

$$\alpha_{ab} = \alpha_a - \alpha_b \tag{3.61}$$

and the net voltage of the junction is:

$$dV_{AB} = \alpha_{AB} dT \tag{3.62}$$

This equation can be used to determine a differential coefficient:

$$\alpha_{AB} = \frac{dV_{AB}}{dT} \tag{3.63}$$

α_{AB} is called the sensitivity of a thermocouple junction because it does not depend on the nature of the junction, as metals may be pressed together, welded, fused, etc, **being not a resistive effect**. What counts is the temperature of the junction and the actual metals. The Seebeck effect is a direct conversion of thermal energy into electric energy.

3.4 PHYSICAL PRINCIPLES OF SENSING

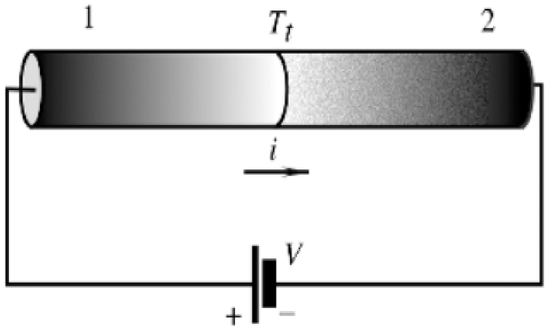
Thermoelectric effect – Peltier effect

The **Peltier effect concerns the reversible absorption of heat which usually takes place when an electric current crosses a junction** between two dissimilar metals. It can be observed as the **opposite effect respect to Seebeck effect**.

The absorption or generation of heat Q_P is a function of the direction of the current:

$$dQ_P = \pm p \cdot i \cdot dt \tag{3.64}$$

where i is the current and t is the time. The coefficient p has the voltage-dimension and it depends on the thermoelectric properties of the material. **It should be noted that heat does not depend on temperature at the junction**.



3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Peltier effect

The effect takes place whether the current is introduced externally or is induced by the thermocouple junction itself (due to the Seebeck effect).

The Peltier effect is used for two purposes: It can **produce heat** or “**produce” cold**, **depending on the direction of electric current through the junction**. This makes it quite useful for the devices where precision thermal control is required.

It should be well understood that the Peltier heat is different from that of the Joule. The Peltier heat depends linearly on the magnitude of the current flow as contrasted to Joule heat ($P = i^2/R$).

The magnitude and direction of Peltier heat do not depend in any way on the actual nature of the contact. It is purely a function of two different bulk materials which have been brought together to form the junction and each material makes its own contribution depending on its thermoelectric properties.

3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Peltier effect

The Peltier effect is a basis for operation of thermoelectric coolers, which are used for the cooling of photon detectors or to stabilize the temperature of semiconductor lasers.

Keeping electronic components at a stable temperature is critical to ensuring that instruments, such as high-power lasers, photodetectors, or camera can properly work. In some cases, cooling below ambient temperature may be necessary.

Passive cooling, based on forced air cooling with a heat sink, can be not enough because:

i) the response to changes in heat load can be slow and inaccurate;

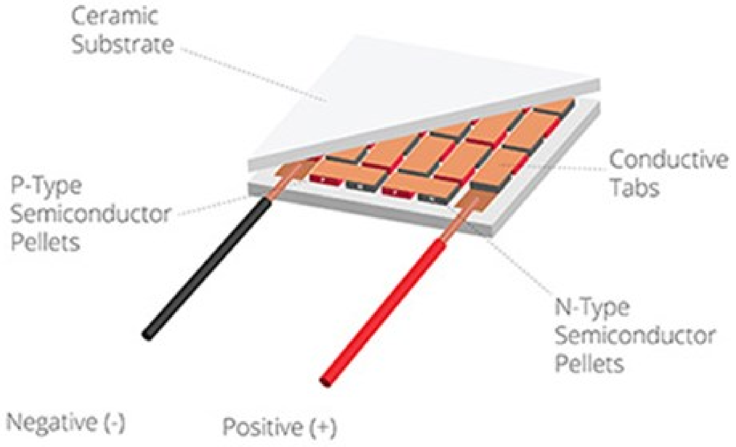
ii) cooling is based on a thermal gradient in which the temperature of the heat source is higher than the ambient temperature.

3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Peltier effect

As an alternative to commonly-used passive cooling techniques, thermoelectric cooling can offer numerous advantages. These include accurate temperature control and faster response, fan-less operation (based on heat sink performance), reduced-noise, space-saving, the ability to cool below ambient temperatures.

The internal structure of a **Peltier cell typically consists of a matrix of n- and p-type semiconductors**. The matrix is electrically connected in series but is thermally arranged in parallel to maximize heat transfer between the hot and cold ceramic surfaces of the module.



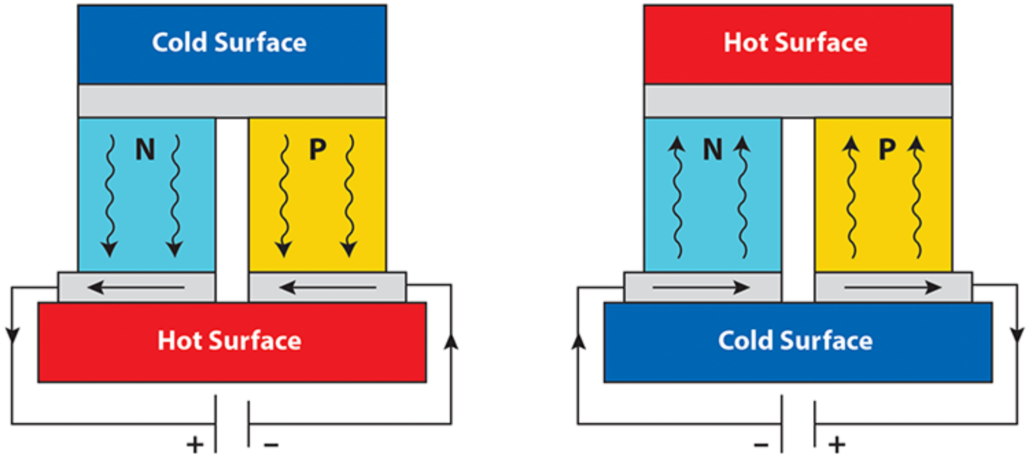
3.4 PHYSICAL PRINCIPLES OF SENSING

Thermoelectric effect – Peltier effect

By Peltier effect, the heat is absorbed or emitted between the junctions of two different conductors, when a current is applied.

In addition, the direction in which heat is transferred can be reversed by simply reversing the direction of current flow.

The current flow causes a heat transfer from one substrate to the other on the opposite side. As a result, the surface on which energy is absorbed becomes cold, and the opposite surface, where energy is released, becomes hot.



3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

Before including a device on a catalogue, a manufacturer must calibrate the device in order to determine the relationship between the device output signal and the stimulus.

Therefore, the calibration operation allows the determination of the transfer function.

Once the transfer function has been determined, all other characteristics can be easily estimated.

Suppose we want to calibrate a gas sensor. The gas sensor can be view as a closed box: **the stimulus represents the gas whose concentration is to be measured, while the response is given by the electrical output signal.**

Gas mixture with known concentration of methane (stimulus)

Methane Sensor

Electrical signal of the photodetector (response)

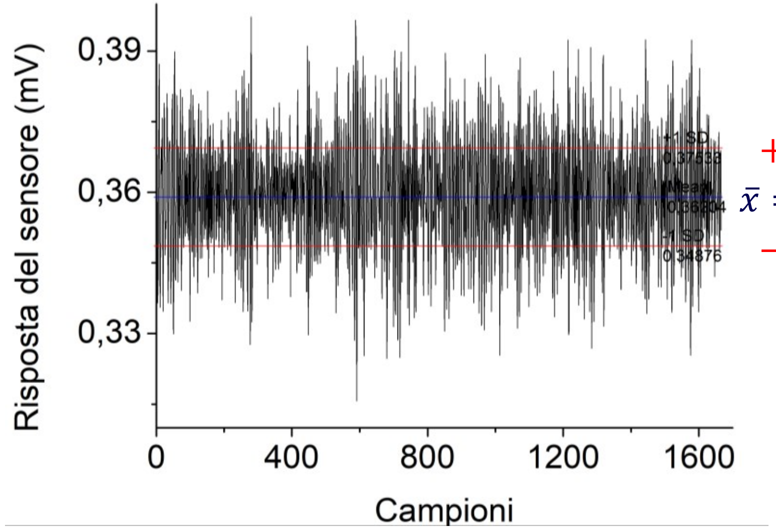
3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

To carry out the calibration it is necessary to use gaseous mixtures with known methane concentrations and to evaluate the sensor response for each of these mixtures.

The sensor response will be related to the concentration of methane in the mixture. We assume error-free concentrations and that the coupling between stimulus and sensor response is instantaneous (no dynamic characteristics). For each methane concentration, N samples were acquired.

The sensor response is the average value of the distribution, while the error was assumed as the 1σ standard deviation of N samples acquired under the same conditions, at a fixed concentration.



$+0.013\ mV$

$\bar{x} = 0.359\ \text{mV}$

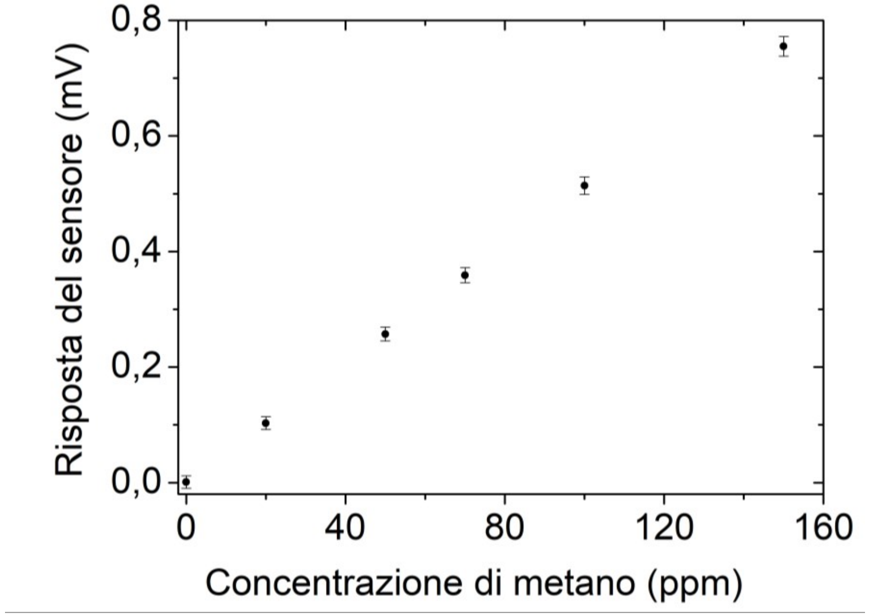
$-0.013\ mV$

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

For each known concentration, the average signal and the standard deviation are calculated.

ppm CH_4 in N_2	Signal (mV)	1σ noise (mV)
0	0.001	±0.011
20	0.103	±0.011
50	0.257	±0.012
70	0.359	±0.013
100	0.514	±0.015
150	0.755	±0.017



The experimental data were represented in the graph, with the relative error bars. The data show a linear trend, thus the Ordinary Least Squares Regression Method can be used to retrieve the sensor transfer function.

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

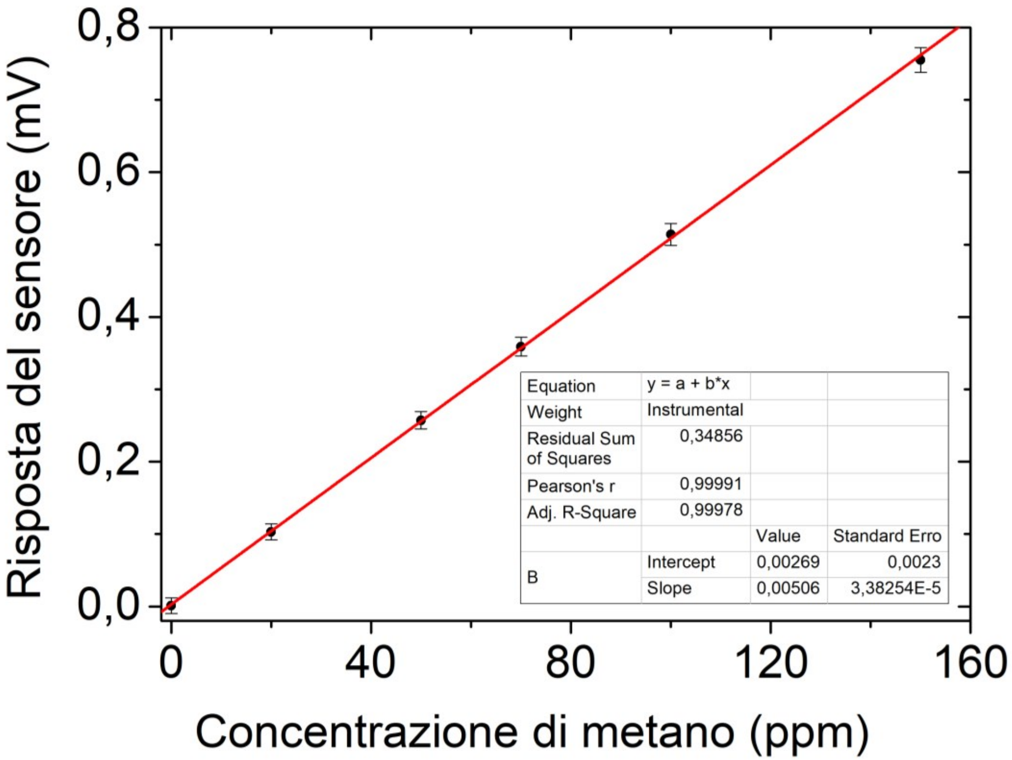
The transfer function is the best linear fit:

$$S[mV] = 0.0027\,mV + 0.00506 \cdot c_{CH_4}\,[ppm]$$

where c_{CH_4} is the methane concentration in the mixture.

The linear fit returns two parameters:

- **Sensitivity** = $0.00506 \pm 0.00003\,\frac{mV}{ppm}$
- **Offset** = $0.0027 \ \pm 0.0023\,mV$



Equation	y = a + b*x		
Weight	Instrumental		
Residual Sum of Squares	0,34856		
Pearson's r	0,99991		
Adj. R-Square	0,99978		
		Value	Standard Erro
B	Intercept	0,00269	0,0023
	Slope	0,00506	3,38254E-5

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

- **Dynamic Range =** 0–150 ppm (the sensor has not been tested with higher concentrations, so we do not know if the sensor responds linearly, and with the same transfer function, even for concentrations> 150 ppm).

- **Full scale output** = 0 mV – 0.755 mV

It is possible to define the sensor's deviation as the difference between the value which is computed from the output voltage and the actual input value.

To convert the measured voltage signal back to the input value, we use the transfer function:

$$c_{CH_4}\ [ppm] = \frac{S[mV] - 0.0027\ mV}{0.00506}$$

Real value (ppm)	Estimated value (ppm)	Deviation (ppm)
0	0.33	-0.33
20	19.82	-0.18
50	50.25	0.25
70	70.42	0.42
100	101.05	1.05
150	148.67	-1.33

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

- **Inaccuracy** represents the highest deviation of a sensor's input value from its true (ideal) value. Let's represent it in the three forms already mentioned:
    - ○ **Directly in terms of measured value**: $\Delta = -1.33\ ppm$
    - ○ **In percent of input span (full scale)**: $\frac{1.33\ ppm}{150\ ppm} x100 \sim 0.9\%$
    - ○ **In terms of output signal**: $0.0027\ mV + 0.00506 \cdot (-1.33\ ppm) = -0.00403\ mV$
- **Hysteresis** can be determined performing the calibration starting first from the lowest concentrations up to the highest ones, and then repeating the procedure in reverse. In this way, it can be established if the sensor suffers from accumulation of residual gas that cannot be expelled.
- **No saturation** was observed in the dynamic range investigated. To determine it, the sensor response should be investigated at higher methane concentrations, in order to find the threshold concentration at which the response function is no longer linear, but flat.

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

- To evaluate any **repeatability errors**, calibration should be performed several times, always under the same experimental conditions (repeatability should not be confused with the 1σ error).

- **Resolution** describes the smallest increments of stimulus which can be sensed. It is represented by the magnitude of the input variation which results in the output smallest step.

  In this case, the smallest step at the output sensor is the last digit that can be visualized on the photodetector, namely 0.001 mV. Then, using the transfer function the resolution R of the sensor can be determined:

$$R[ppm] = \frac{0.001\ mV - 0.0027\ mV}{0.00506} = 0.33\ ppm$$

3.5 EXAMPLE: SENSOR CALIBRATION

Calibration of a gas sensor

- **Signal-to-noise ratio (SNR)** is a measure used in science and engineering that compares the level of a desired signal to the level of background noise. It is defined as the ratio of signal power to the noise power, sometimes expressed in decibels: $SNR = \frac{Signal}{Noise}$
- **Limit of detection (LoD) or Minimum Detection Limit (MDL) or Noise Equivalent Concentration (NEC)** is the lowest analyte concentration at which detection is feasible. In gas sensing, LoD is typically determined as the concentration providing a SNR = 1.
- **Limit of quantification (LoQ)** is the lowest concentration at which the analyte can not only be reliably detected but at which some predefined goals for bias and imprecision are met. The LoQ may be equivalent to the LoD or it could be at a much higher concentration (typically 5-10 LoD).