Performance Characteristics of Sensors and Actuators

The performance characteristics of the device or system are the most important issues the engineer is faced with? Why

- If we need to sense temperature, then of course a temperature sensor is needed.
- But what kind of sensor, what temperature range can it sense?
- How "accurate" does it need to be?
- Is it important that it be a linear measurement, and how critical is the repeatability of the sensor?
- Does it need to respond quickly or can we use a slow responding sensor?

Input and Output

- Sensors

Input: stimulus or measurand (temperature, pressure, light intensity, etc.)

Ouput: electrical signal (voltage, current, frequency, phase, etc.)

Variations: output can sometimes be displacement (thermometers, magnetostrictive and piezoelectric sensors). Some sensors combine sensing and actuation

Input and Output

- Actuators

Input: electrical signal (voltage, current, frequency, phase, etc.)

Output: mechanical (force, pressure, displacement) or display function (dial gauge, light indication, display, etc.)

In addition, we must take into account input and output properties such as impedance, temperature, and environmental conditions in order to provide proper operating conditions for the device.

Transfer function

- The characteristics of a device start with its transfer function, that is, the relation between its input and output. This includes many other properties, such as span (or range), frequency response, accuracy, repeatability, sensitivity, linearity, reliability, and resolution, among others
- Relation between input and output
- Other names:
  - Input output characteristic function
  - transfer characteristic function
  - Response of a device
- usually defined by some kind of mathematical equation and a descriptive curve or graphical representation in a given range of inputs and outputs.

Transfer function (cont.)

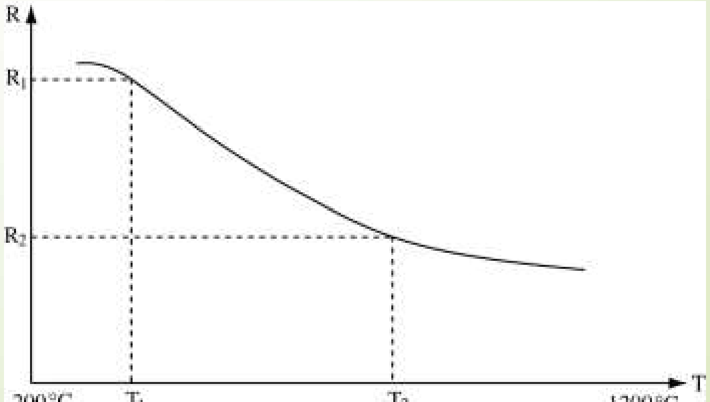
- Linear or nonlinear
- Single valued or not
- One dimensional or multi dimensional
  - Single input, single output
  - Multiple inputs, single output
- In most cases:
  - Difficult to describe mathematically (given graphically)
  - Often must be defined from calibration data
  - Often only defined on a portion of the range of the device

Transfer function (cont.)

- T_1 to T_2 - approximately linear
  - Most useful range
  - Typically a small portion of the range
  - Often taken as linear

where x is the input (stimulus in sensors or, say, current to an actuator) and S is the output. The dependence of the output S on x indicates that this function can be (and often is) nonlinear

$$\boxed{S = f(x)}$$



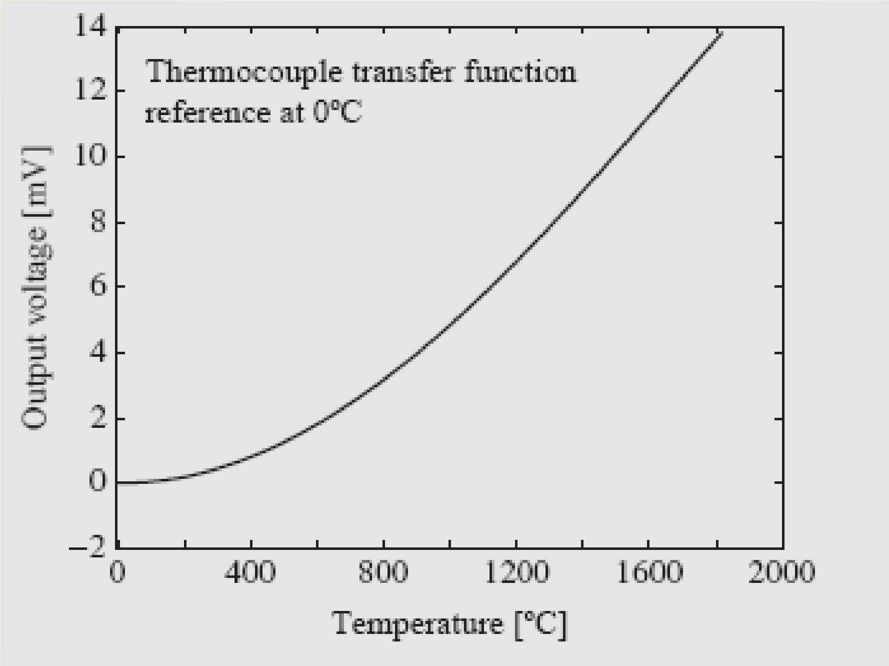
Thermocouple transfer functions high-order polynomials

The output (voltage) of a thermocouple (temperature sensor) for a given temperature is given by a polynomial that can range from a 3rd order to a 12th order polynomial depending on the type of thermocouple. The output of a particular type of thermocouple is given by the following relation in the range 0°C–1820°C:

$$\begin{aligned} V = (&-2.4674601620 \times 10^{-1} \times T + 5.9102111169 \times 10^{-3} \times T^2 \\ &- 1.4307123430 \times 10^{-6} \times T^3 + 2.1509149750 \times 10^{-9} \times T^4 \\ &- 3.1757800720 \times 10^{-12} \times T^5 + 2.4010367459 \times 10^{-15} \times T^6 \\ &- 9.0928148159 \times 10^{-19} \times T^7 + 1.3299505137 \times 10^{-22} \times T^8) \times 10^{-3}\ \mathrm{mV}. \end{aligned}$$

This is a rather involved transfer function (most sensors will have a much simpler response) and is nonlinear. The main purpose of the elaborate function is to provide very accurate representation over the range of the sensor (in this case 0 C–1820 C).

Thermocouple transfer functions high-order polynomials



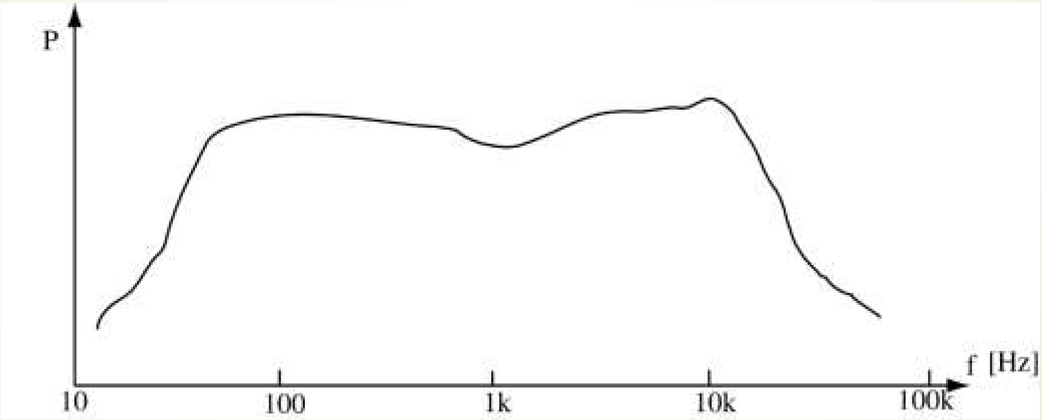
For example, the output of a thermocouple is typically 10–50 μV/C, whereas a piezoelectric sensor can produce 300 V or more in response to motion. A magnetic actuator may require, say, 20 A at 12 V, whereas an electrostatic actuator may operate at 500 V (or higher) at a very low current.

Transfer function (cont.)

- Other data from transfer function
  - saturation
  - sensitivity
  - full scale range (input and output)
  - hysteresis
  - Deadband
  - etc.

Transfer function (cont.)

- Other types of transfer functions
  - Response with respect to a given quantity
  - Performance characteristics (reliability curves, etc.)
  - Viewed as the relation between any two characteristics



Impedance and impedance matching

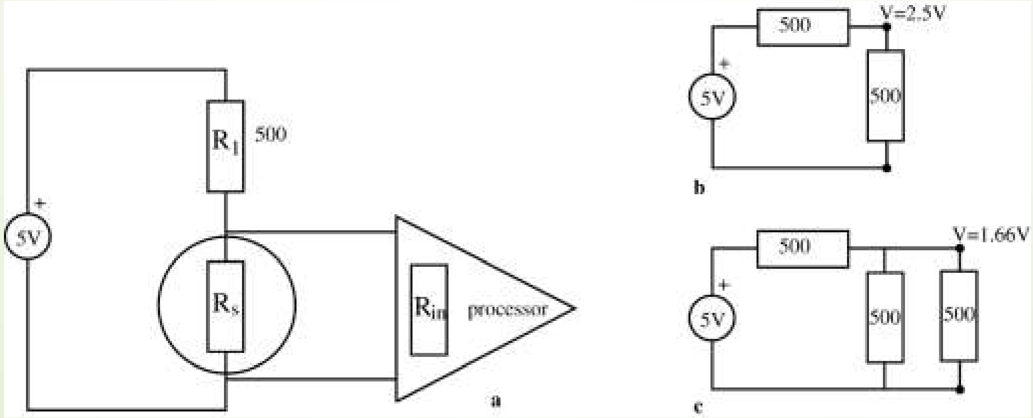
- Input impedance: ratio of the rated input voltage and the resulting current through the input port of the device with the output port open (no load)
- Output impedance: ratio of the rated output voltage and short circuit current of the port (i.e. current when the output is shorted)
- These are definitions for two-port devices
- The reason these properties matter is that they affect the operation of the device.

Impedance (cont.)

- Sensors: only output impedance is relevant
- Actuators: only input impedance is relevant
- Can also define mechanical impedance
  - Not needed - impedance is important for interfacing
  - Will only talk about electrical impedance

Impedance (cont.)

- Why is it important? It affects performance
- Example: 500 W sensor (output impedance) connected to a processor
  - b. Processor input impedance is infinite
  - c. Processor input impedance is 500 W



Impedance (cont.)

- Example. Strain gauge: impedance is 500 Ω at zero strain, 750 Ω at measured strain
- b: sensor output: 2.5V (at zero strain), 3V at measured strain
- c. sensor output: 1.666V to 1.875V
- Result:
  - Loading in case c.
  - Reduced sensitivity (smaller output change for the same strain input)
  - b. is better than c (in this case). Infinite impedance is best.

Impedance (cont.)

- Current sensors: impedance is low - need low impedance at processor
- Same considerations for actuators
- Impedance matching:
  - Sometimes can be done directly (most devices have very high input impedances)
  - Often need a matching circuit
  - From high to low or from low to high impedances

Impedance (cont.)

- Impedance can (and often is) complex: $Z=R+jX$
- In addition to the previous:
  - Conjugate matching ($Z_{in}=Z_{out}^*$) - maximum power transfer
    - Critical for actuators!
    - Usually not important for sensors
    - (input impedance actuator) $Z_{in}=R+jX$,
  - No reflection matching ($Z_{in}=Z_{out}$) - no reflection from load
    - Important at high frequencies (transmission lines)
    - Equally important for sensors and actuators (antennas)

➨In the case of no reflection of current and voltage the impedance of the sensor or actuator must equal the input impedance of the processor. This requirement does not guarantee maximum power transfer, only nonreflection.

The following considerations are relevant in component interconnection:

- Characteristics of the interconnected components (e.g., domain of the component [mechanical, electrical/electronic, thermal], type of the component [actuator, sensor, drive circuit, controller, mounting or housing])
- Purpose of the interconnected system (e.g., drive a load, measure a signal, communicate information, minimize noise and disturbances—mechanical shock and vibration in particular).
- Signal/power levels of operation.

Examples of across variables are voltage, velocity, temperature, and pressure. Examples of through variables are current, force, heat transfer rate, and fluid flow rate.

Several objective categories of impedance matching are given in the following:

- **Source and load matching for maximum power transfer:** In a drive system, an important objective may be to maximize the power transmitted from the power source to the actuator or the load. [Proper impedance matching can achieve the requirement of maximum power transfer].
- **Power transfer at maximum efficiency:** Achieving maximum efficiency in power transfer is different from achieving maximum power simply because maximum efficiency is not achieved at maximum power transfer. [The load impedance can be properly chosen to achieve high efficiency].
- **Reflection prevention in signal transmission:** When two components are connected by a cable (e.g., coaxial cable) with characteristic impedance (e.g., 50 or 75 Ω for a coaxial cable), due to the impedance difference at the two ends (due to the impedances of the connected components), there will be signal reflection (similar to elastic wave reflection due to density difference in two media).

[The end impedances and the characteristic impedance of the cable have to be matched in order to avoid signal reflection].

- **Loading reduction:** When two components are interconnected, in some applications, it is required that the output component does not load the input component. For example, in a sensing process, the sensor should not alter the conditions of the sensed object. In other words, the measuring instrument should not distort the signal that is measured.
- [An impedance transformer would be required to achieve proper impedance matching for loading reduction].

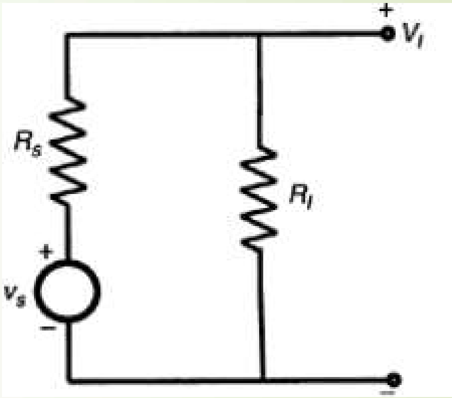
A device with a ***high-input impedance*** has the further advantages that it will consume less power (i.e., v2/R is low) particularly from the input device to which it is connected, for a given input voltage, and furthermore, the power transfer will take place at higher efficiency. The fact that an output device having ***low input impedance*** extracts a high level of power from its input device may also be interpreted as the reason for loading error in the input device.

Impedance Matching

Consider a dc power supply of voltage u*s* and output impedance (resistance) *Rs*. It is used to power a load of resistance R_l, as shown in following figure. What should be the relationship between R_s and R_l if the objective is to maximize the power absorbed by the load?

- Solution:
- Current through the circuit is

$$i_l = \frac{v_s}{R_l + R_s}$$



- the voltage across the load is

$$v_l = i_l R_l = \frac{v_s R_l}{R_l + R_s}$$

- The power absorbed by the load is
- For maximum power, we need

$$p_l = i_l v_l = \frac{v_s{}^2 R_l}{[R_l + R_s]^2}$$

$$\frac{dp_l}{dR_l} = 0$$

- This gives the requirement for maximum power as $R_l = R_s$

Impedance Matching

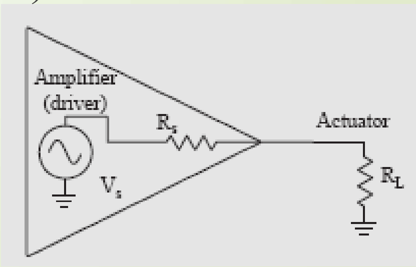
Example: A voice coil actuator (an actuator that operates on the principle of the loudspeaker; it is pulse driven by an amplifier. The amplifier provides an amplitude $V_s = 12$ V. The internal impedance of the amplifier is $R_s = 4\ \Omega$.

a. Calculate the power transferred to an impedance-matched actuator.

b. Show that the power transmitted to an actuator with lower or higher impedance is lower than that for the matched actuator in (a).

c. What is the power supplied to a 4 W actuator if the internal impedance of the amplifier is $0.5\ \Omega$ and supplies the same voltage (12 V)?

- **Solution:** Because the actuator is pulse driven, the power is considered instantaneous, but since the voltage is constant through the duration of the pulse, we will calculate the power as if it were a DC source (i.e., the power during the ON portion of the pulse).



Impedance Matching

a. The equivalent circuit for the matched condition is shown in the above figure. The actuator's resistance is R_L = 4 W and the power supplied to the actuator is

$$P_L = \frac{V_L^2}{R_L} = \left(\frac{V_s}{R_s + R_L} R_L\right)^2 \frac{1}{R_L} = \left(\frac{12}{4+4} 4\right)^2 \frac{1}{4} = 9 \text{ W}.$$

Note that this is exactly half the power supplied by the source, the other half being dissipated on the internal resistance of the source in the form of heat.

b. With a lower or higher actuator impedance we have, for R_L= 2 Ω for example,

$$P_L = \frac{V_L^2}{R_L} = \left(\frac{V_s}{R_s + R_L} R_L\right)^2 \frac{1}{R_L} = \left(\frac{12}{4+2} 2\right)^2 \frac{1}{2} = 8 \text{ W}.$$

Similarly, for a higher actuator impedance, say for R_L = 6 Ω,

$$P_L = \frac{V_L^2}{R_L} = \left(\frac{V_s}{R_s + R_L} R_L\right)^2 \frac{1}{R_L} = \left(\frac{12}{4+6} 6\right)^2 \frac{1}{6} = 8.64 \text{ W}.$$

Clearly, maximum power is transferred for matched impedances.

Range and Span

- **Range:** lowest and highest values of the stimulus
- **Span:** the arithmetic difference between the highest and lowest values of the stimulus that can be sensed within acceptable errors (the difference between the range values)
- **Input full scale** (IFS) = span
- **Output full scale** (OFS): difference between the upper and lower ranges of the output of the sensor corresponding to the span of the sensor
- **Dynamic range:** ratio between the upper and lower limits and is usually expressed in dB

Range and Span (Cont)

- Example: a sensors is designed for: -30 °C to +80 °C to output 2.5V to 1.2V
- Range: -30°C and +80 °C
- Span: 80- (-30)=110 °C
- Input full scale = 110 °C
- Output full scale = 2.5V-1.2V=1.3V
- Dynamic range=20log(110/30)=11.29 dB
- Dynamic range is the span where the minimum to maximum strength signals can be detected and measured before unwanted artifacts appear above the noise floor.

Range, Span and Dynamic Range (Cont.)

- The ratio of the dynamic range represents either powerlike (power, power density) or voltagelike quantities (voltage, current, force, fields, etc.), the dynamic range is written as follows:
- For voltage-like quantities:

$$\text{Dynamic range} = 20 \log|\text{span}/\text{lower measurable quantity}|.$$

- For power-like quantities:

$$\text{Dynamic range} = 10 \log|\text{span}/\text{lower measurable quantity}|$$

- Suppose we look at a 4-digit digital voltmeter capable of measuring between 0 and 20 V. The total span is 19.99 V and the resolution (smallest increment) is 0.01 V. The dynamic range is thus

$$\text{Dynamic range} = 20 \log(19.99/0.01) = 20 \times 3.3 = 66 \text{ dB}.$$

- On the other hand, a 4-digit digital wattmeter measuring up to 20 W in increments of 0.01 W will have a dynamic range of

$$\text{Dynamic range} = 10 \log(19.99/0.01) = 10 \times 3.3 = 33 \text{ dB}.$$

Range and Span (cont.)

- Range, span, full scale and dynamic range may be applied to actuators in the same way
- Span and full scale may also be given in dB when the scale is large.
- In actuators, there are other properties that come into play:
  - Maximum force, torque, displacement
  - Acceleration
  - Time response, delays, etc.

Range and Span (cont.)

- Example: A silicon temperature sensor has a range between 0 °C and 90°C. The accuracy is defined in the data sheet as 0.5°C. Calculate the dynamic range of the sensor.
- Solution: The resolution is not given, so we will take the accuracy as the minimum measurable quantity. In general, these need not be the same. Since the minimum resolution is 0.5, the dynamic range is

$$\text{Dynamic range} = 20\log_{10}\left(\frac{90}{0.5}\right) = 45.1 \text{ dB.}$$

- Example: A loudspeaker is rated at 6 W and requires a minimum power of 0.001 W to overcome internal friction. What is its dynamic range? Clearly, any change smaller than 1 mW will not change the position of the speaker's cone and hence no change in output will be produced. Thus 1 mW is the resolution of the speaker and the dynamic range of the speaker is

$$\text{Dynamic range} = 10\log_{10}\left(\frac{6}{0.001}\right) = 37.78 \text{ dB.}$$

Range and Span (cont.)

- Example: A 16-bit analog to digital [A/D] converter is used to convert an analog music recording into digital format so it can be stored digitally and played back (by converting it back to analog form). The amplitude varies between - 6 V and 6 V.

a. Calculate the smallest signal increment that can be used.

b. Calculate the dynamic range of the A/D conversion.

- Solution:

a. A 16-bit A/D converter can represent 2^{16} = 65,536 levels of the signal. The smallest signal increment in the case discussed here is

$$\Delta V = \frac{12}{2^{16}} = 1.831 \times 10^{-4}\ \text{V}.$$

That is, the signal is represented in increments of 0.1831 mV.

b. The dynamic range of the A/D converter is

$$\textit{Dynamic range} = 6.0206N = 96.33\ \text{dB}.$$

Accuracy, errors, repeatability

- **Errors:** deviation of the device output from the "ideal"
- In accuracy sources in the output (i.e. from transfer function):
  - materials used
  - construction tolerances
  - ageing
  - operational errors
  - calibration errors
  - matching (impedance) or loading errors
  - noise
  - many others

Accuracy, errors (cont.)

- **Errors:** defined as follows:
- a. As a difference: $e = V - V_0$ (V_0 is the actual value, V is that measured value by the device (the stimulus in the case of sensors or output in actuators).
- b. As a percentage of internal full scale (input) or (span for example), $e = \Delta t / (t_{max} - t_{min}) * 100$ where t_{max} and t_{min} are the maximum and minimum values at which the device is designed to operate (range values).
- c. In terms of the output signal expected rather than the stimulus. Or the difference between values or it may be represented as a percentage of OFS.

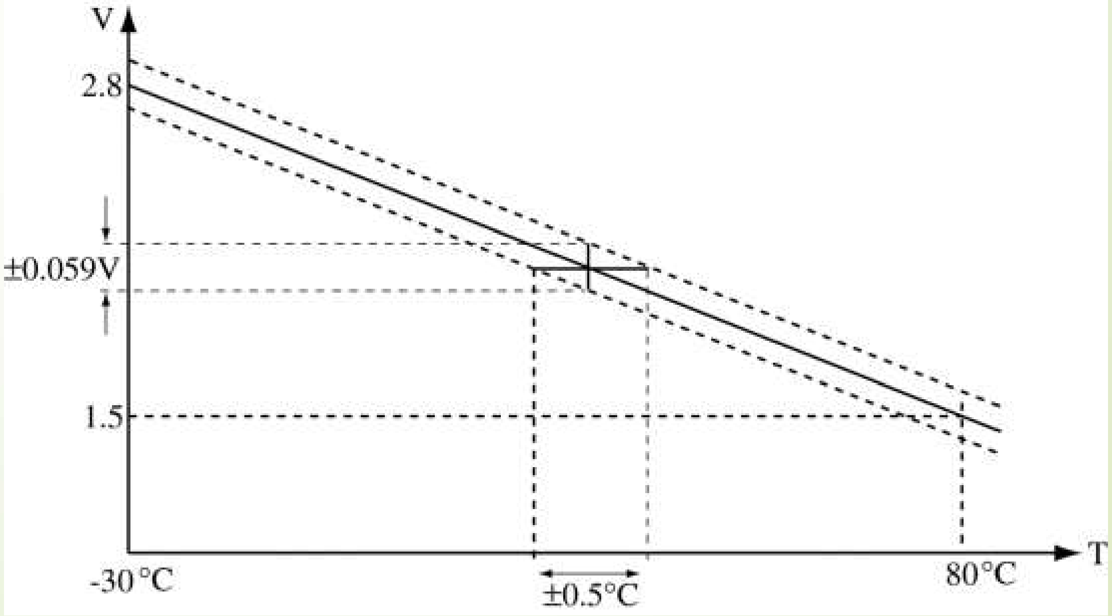
Example: errors

- **Example:** A thermistor is used to measure temperature between –30 and +80 °C and produce an output voltage between 2.8V and 1.5V. Because of errors, the accuracy in sensing is ±0.5°C.
- The **accuracy** of a measurement or approximation is the degree of closeness to the exact value. The **error** is the difference between the approximation and the exact value.

Example (cont)

- In terms of the input as ± 0.5°C
- Percentage of input: e = 0.5/(80+30)*100 = 0.454%
- In terms of output. From the transfer function: e = ±0.059V.

In most cases, the error given in terms of the measurand or percent of IFS is the best measure of the sensor's accuracy.



More on errors

- **Static errors**: not time dependent
- **Dynamic errors**: time dependent
- **Random errors**: Different errors in a parameter or at different operating times, (e.g. the thermal noise of a sensor). An error is considered random if the value of what is being measured sometimes goes up or sometimes goes down. A very simple example is **our blood pressure**. Even if someone is healthy, it is normal that their blood pressure does not remain exactly the same every time it is measured.
- **Systemic errors**: errors are constant at all times and conditions. There are four types of systematic error: observational, instrumental, environmental, and theoretical.

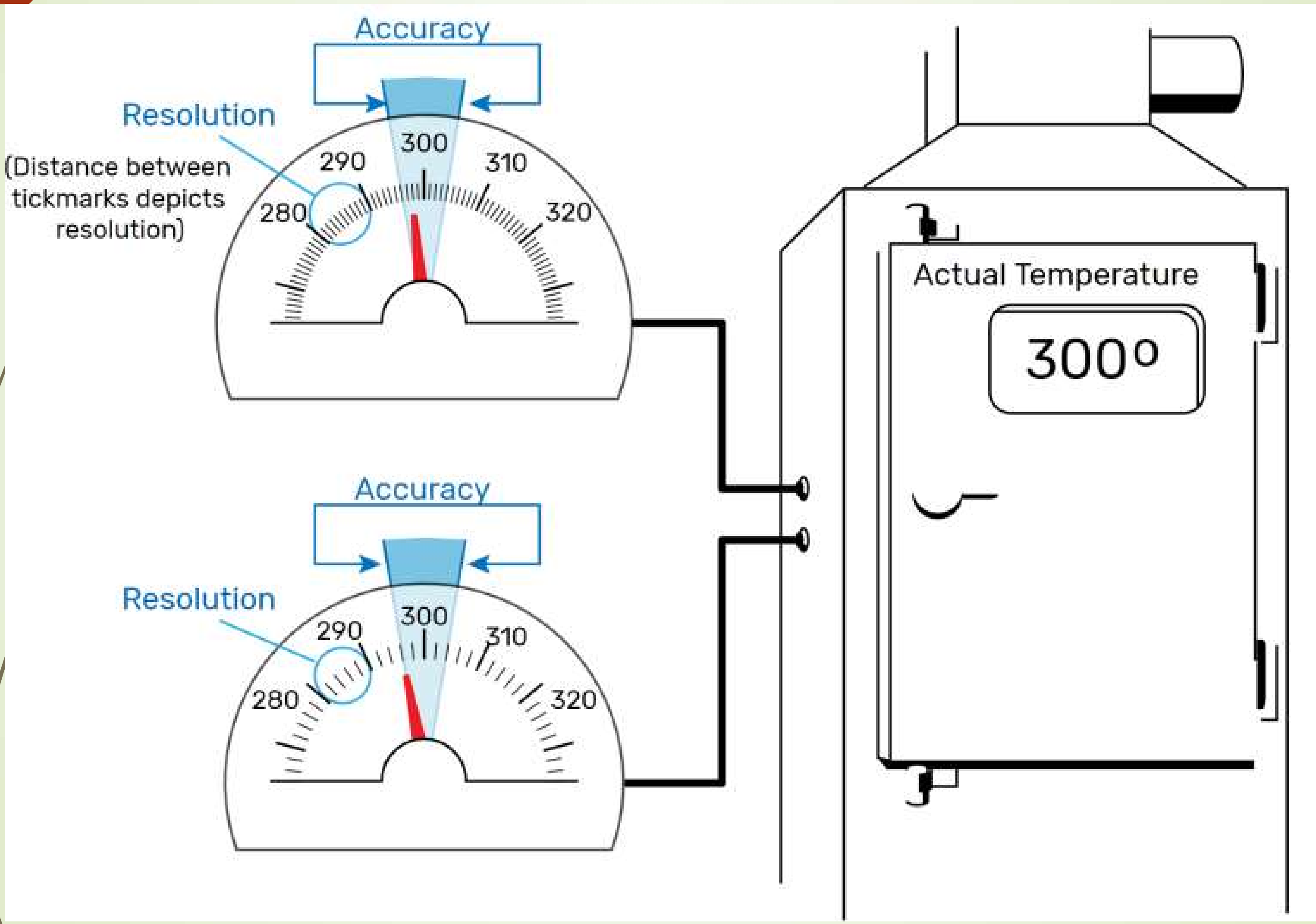
Accuracy and Resolution

What's the difference between accuracy and resolution? Accuracy is how close a reported measurement is to the true value being measured. Resolution is the smallest change that can be measured.

The accuracy of these temperature gauges is +/-4 degrees, meaning they can be different from the correct value by four degrees in either direction. (Typically, accuracy specifications are stated as a plus-minus range.)

The gauge on top has finer resolution. Notice that there are more tick marks between 280 and 290 on the top gauge than on the bottom one. Finer resolution reduces rounding errors, but doesn't change a device's accuracy. However, resolution that is too coarse may add rounding errors.

Example of Resolution and accuracy



Repeatability

- Also called reproducibility: failure of the sensor or actuator to represent the same value (i.e. stimulus or input) under identical conditions when measured at different times.
  - usually associated with calibration
  - viewed as an error.
  - given as the maximum difference between two readings taken at different times under identical input conditions.
  - error given as percentage of input full scale.

Sensitivity

- Sensitivity of a sensor is defined as the change in output for a given change in input, usually a unit change in input. Sensitivity represents the slope of the transfer function. Or, more generally, the *minimum input of physical parameter that will create a detectable output change*
- Same for actuators

$$s = \frac{dS}{dx} = \frac{d}{dx}(f(x)).$$

- For example, a typical blood pressure transducer may have a sensitivity rating of 10 mV/V/mm Hg; that is, there will be a 10-mV output voltage for each volt of excitation potential and each mm Hg of applied pressure.

Sensitivity (cont.)

- Example for a linear transfer function:
  - Note the units
  - a is the slope
- For the transfer function, where the output is resistance (R) and the input is temperature (T), we have

$$s = \frac{dR}{dT} = \frac{d}{dT}(aT + b) = a \quad [\Omega/^{\circ}\mathrm{C}].$$

➡Usually sensitivity is associated with sensors. However, as long as a transfer function can be defined for an actuator, the same ideas can be extended to actuators. Therefore it is quite appropriate to define a sensitivity of, say, a speaker as *dP / dI*, where *P* is the pressure (output) the speaker generates per unit current I (input) into the speaker, or the sensitivity of a linear positioner as *dl / dV*, where *l* is linear distance and *V* is the voltage (input) to the positioner.

Sensitivity (cont.)

- Usually associated with sensors
- Applies equally well to actuators
- Can be highly nonlinear along the transfer function
- Measured in units of output quantity per units of input quantity (W/°C, N/V, V/°C, etc.)

Sensitivity (cont.)

- Example: The response of a pressure sensor is determined experimentally and given in the table below. A noise in the form of pressure variations of 330 Pa exists due to local atmospheric changes. Calculate the output due to noise and the error in output produced by this noise.

Pressure [kPa]	100	120	140	160	180	200	220	240	260	280	300	320	340	360	380	400
Voltage [V]	1.15	1.38	1.6	1.86	2.1	2.35	2.6	2.89	3.08	3.32	3.59	3.82	4.05	4.29	4.54	4.78

- ***Solution:*** to calculate the output due to noise we must first find the sensitivity of the sensor. Since the output is experimental, we first need to pass a best fit curve through the data.

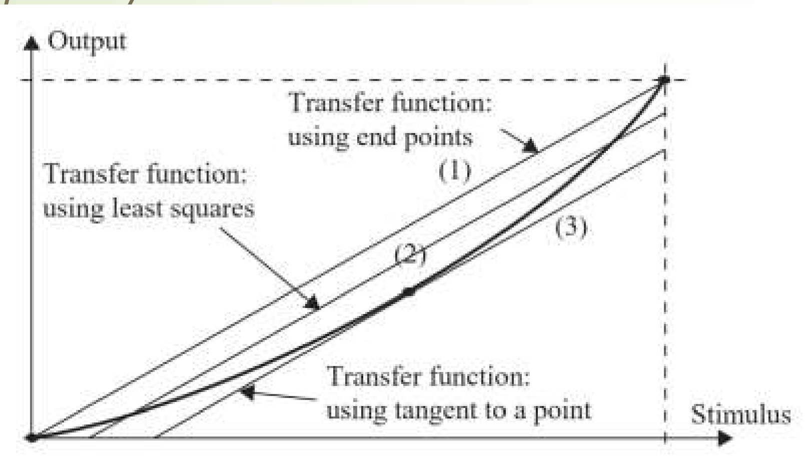
$$V = aP + b = 0.0122P - 0.0783\,\mathrm{V},$$

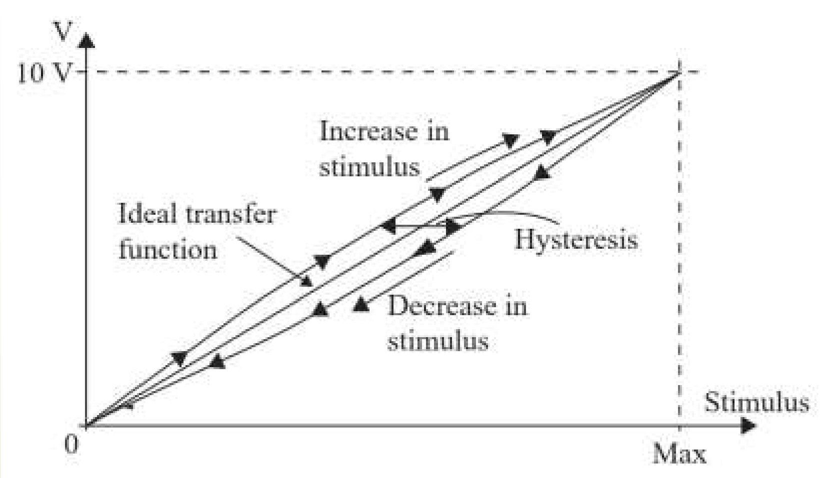
- where P is pressure in kilopascals. Thus the sensitivity is

$$s = \frac{dV}{dP} = a = 0.0122\ \mathrm{V/kPa}.$$

Hysteresis

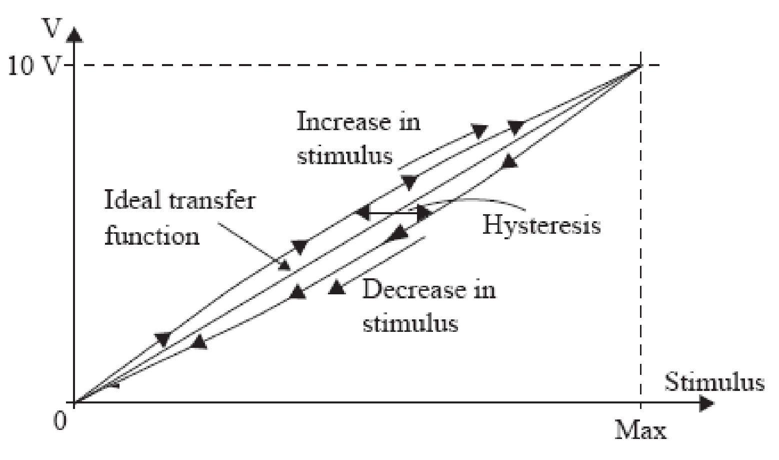
- **Hysteresis** (literally lag) - the deviation of the sensor's output at any given point when approached from two different directions
- Caused by electrical or mechanical systems
  - Magnetization
  - Thermal properties
  - Loose linkages





Hysteresis - Example

- For example: If temperature is measured, at a rated temperature of 50°C, the output might be 4.95 V when temperature increases but 5.05 V when temperature decreases.
- This is an error of ±0.5% (for an output full scale of 10V in this idealized example).
- Hysteresis is also present in actuators and, in the case of motion, more common than in sensors.



Nonlinearity

- A property of the sensor (nonlinear transfer function) or:
- Introduced by errors
- Nonlinearity errors influence accuracy.
- Nonlinearity is defined as the maximum deviation from the ideal linear transfer function.
- The latter is not usually known or useful
- Nonlinearity must be deduced from the actual transfer function or from the calibration curve
- A few methods to do so:

Saturation

- **Saturation** a property of sensors or actuators when they no longer respond to the input.
- Usually at or near the ends of their span and indicates that the output is no longer a function of the input or, more likely is a very nonlinear function of the input.
- Should be avoided - sensitivity is small or nonexistent
- In actuators, it can lead to failure of the actuator (increase in power loss, etc.)

Calibration

- Calibration: the experimental determination of the transfer function of a sensor or actuator.
- Typically, needed when the transfer function is not known or,
- When the device must be operated at tolerances below those specified by the manufacturer.
- Example, use a thermistor with a 5% tolerance on a full scale from 0 to 100°C to measure temperature with accuracy of, say, ±0.5°C.
- The only way this can be done is by first establishing the transfer function of the sensor.

Calibration (cont.)

- Two methods:
- A . known transfer function:
  - Determine the slope and crossing point (line function) from two known stimuli (say two temperatures) if the transfer function is linear
  - Measure the output
  - Calculate the slope and crossing point in $V = a T + b$
  - If the function is more complex, need more points: $V = a T + bT^2 + cT^3 + d$
  - 4 measurements to calculate a, b, c, d
  - Must choose points judiciously - if linear, use points close to the range. If not, use equally spaced points or points around the locations of highest curvature

Calibration (cont.)

- Two methods:
- b. Unknown transfer function:
  - Measure the output R_i at as many input values T_i as is practical
  - Use the entire span
  - Calculate a best linear fit (least squares for example)
  - If the curve is not linear use a polynomial fit
  - May use piecewise linear segments if the number of points is large.

Calibration (cont.)

- Calibration is sometimes an operational requirement (thermocouples, pressure sensors)
- Calibration data is usually supplied by the manufacturer
- Calibration procedures must be included with the design documents
- Errors due to calibration must be evaluated and specified

Resolution

- **Resolution:** the minimum increment in stimulus to which it can respond. It is the magnitude of the input change which results in the smallest discernible output.
- Example: a digital voltmeter with resolution of 0.1V is used to measure the output of a sensor. The change in input (temperature, pressure, etc.) that will provide a change of 0.1V on the voltmeter is the resolution of the sensor/voltmeter system.

Resolution (cont.)

- Resolution is determined by the whole system, not only by the sensor
- The resolution of the sensor may be better than that of the system.
- The sensor itself must interact with a processor, the limiting factor on resolution may be the sensor or the processor.
- Resolution may be specified in the units of the stimulus (0.5°C for a temperature sensor, 1 mT for a magnetic field sensor, 0.1mm for a proximity sensor, etc) or may be specified as a percentage of span (0.1% for example).

Resolution (cont.)

- In digital systems, resolution may be specified in bits (1 bit or 6 bit resolution)
- In analog systems (those that do not digitize the output) the output is continuous and resolution may be said to be infinitesimal or mV (for the sensor or actuator alone).
- Resolution of an actuator is the minimum increment in its output which it can provide.
- Example: a stepper motor may have 180 steps per revolution. Its resolution is 2°.
- A graduated analog voltmeter may be said to have a resolution equal to one graduation (say 0.01V). (higher resolution may be implied by the user who can easily interpolated between two graduations.

Resolution (cont.)

- In digital systems, resolution may be specified in bits (such as N-bit resolution) or in some other means of expressing the idea of resolution. In an analog-to-digital (A/D) converter, the resolution means the number of discrete steps the converter can convert.
- For example, a 12-bit resolution means the device can resolve $2^{12} = 4096$ steps. If the converter digitizes a 5 V input, each step is $5/4096 = 1.22 \times 10^{-3}$ V. On the analog side, the resolution may be described as 1.22 mV, but on the digital side it is described as12-bit resolution.
- In digital cameras and in display monitors the resolution is typically given as the total number of pixels. Thus a digital camera may be said to have a resolution of a number (x) of megapixels.