18ECO133T - SENSORS AND TRANSDUCERS

PREPARED BY : Dr. G.RAMYA
A.P (Sr. G)/ ECE
SRMIST

18ECO133T - SENSORS AND TRANSDUCERS

UNIT 1

- Introduction to sensors/ transducers, principles, classification based on different criteria,
- Characteristics of measurement systems, Static characteristics - Accuracy, Precision, Resolution, Sensitivity,
- Dynamic characteristics and Environmental Parameters,
- Characterization and its type - Electrical characterization,
- Mechanical Characterization, Thermal Characterization,
- Optical Characterization, Errors and its classification,
- Selection of transducers, Introduction to mechanical sensors
- Resistive potentiometer and types, Strain gauge: Theory, type, design consideration, sensitivity.
- Resistive transducer: RTD, materials used in RTD, Thermistor: thermistor material, shape.

Transducer, Sensor, and Actuator

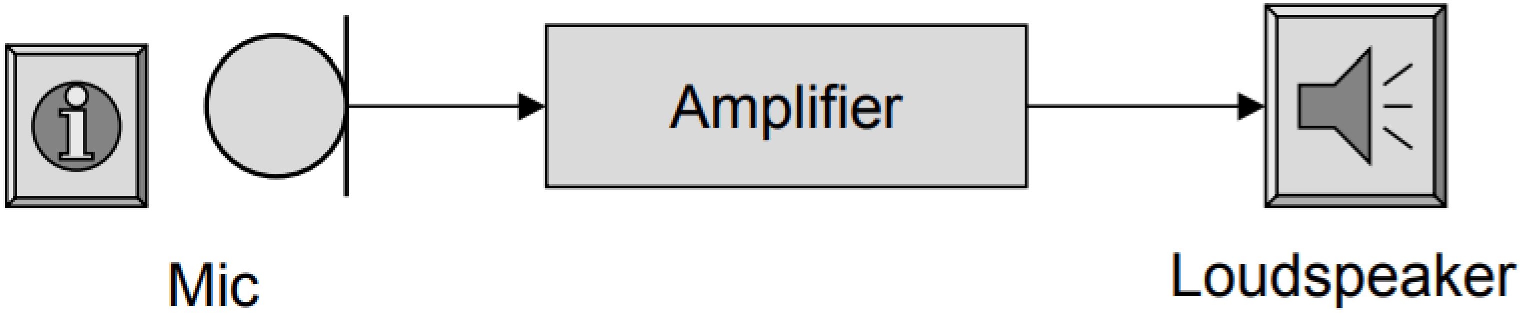
- Transducer:
  - a device that converts energy from one form to another
- Sensor:
  - converts a physical parameter to an electrical output (a type of transducer, e.g. a microphone)
- Actuator:
- converts an electrical signal to a physical output (opposite of a sensor, e.g. a speaker)

SENSORS and TRANSDUCERS

SENSORS convert energy information

One energy form must be converted into the same or another energy form with exactly the same information content as the originating energy form

Example:

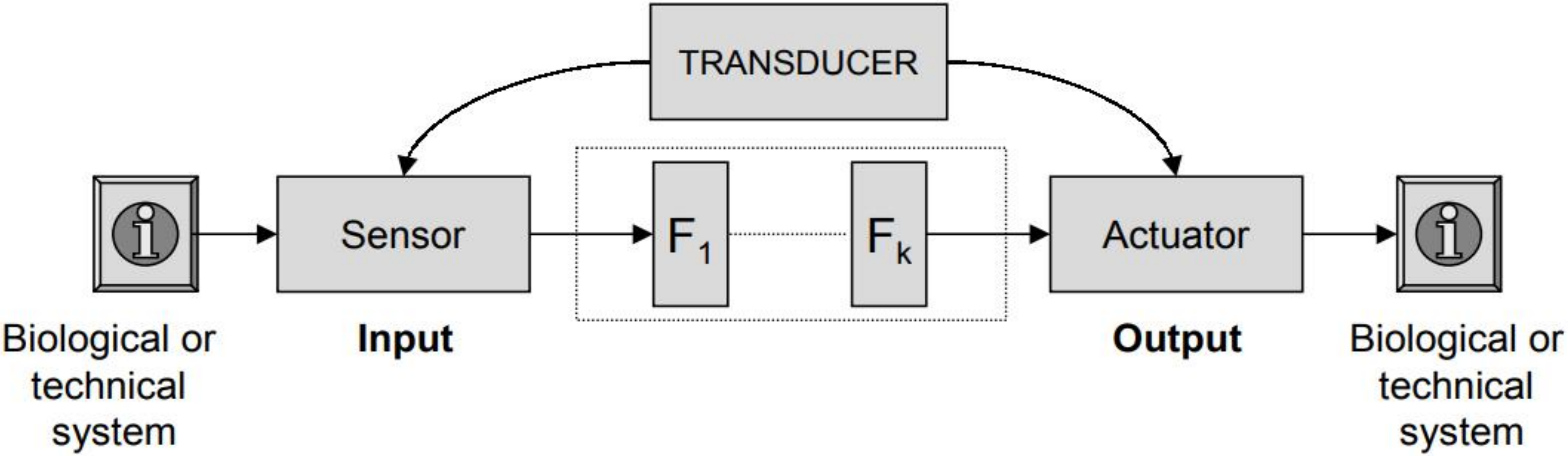


SENSORS and TRANSDUCERS

TRANSDUCER - *latin tranducere* - 'to convert'

Input transducer - sensor

Output transducer - actuator



SESSION 1

- Introduction to sensors/ transducers
- Principles
- Classification based on different criteria
  1. BASED ON TRANSDUCTON PRINCIPLES
  2. TECHNOLOGY BASED CLASSIFICATION
  3. APPLICATION BASED CLASSIFICATION
  4. PROPERTY BASED CLASSIFICATION
  5. BASED ON EMERGING SENSOR TECHNOLOGIES

What are Sensors /Transducers

- Definition from Instrument Society of America
- A sensor or transducer as a device which provides a usable output in response to a specified measurand.

Output is defined as an 'electrical quantity' and

Measurand -Physical quantity, or condition which is measured.’

- Physical quantity :Temperature, Pressure, force, motion, displacement, humidity, light flow etc.
- Electrical quantity: Change in resistance, inductance, capacitance etc.

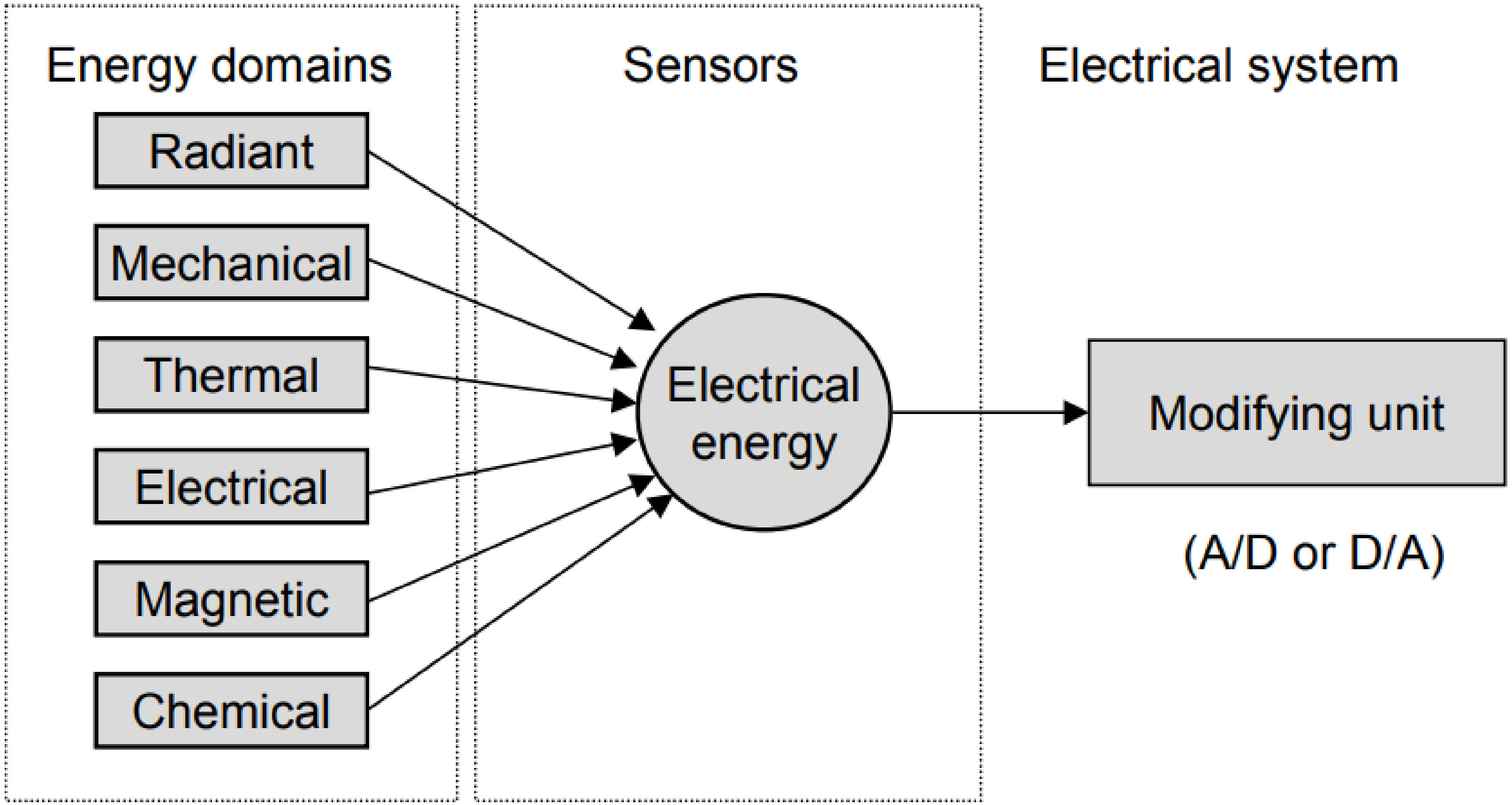
- Generalized Definition is extending 'electrical quantity' to any type of signal such as mechanical and optical and extending 'physical quantity, property, or condition being measured' to those of nature--chemical and biochemical and so on.

PRINCIPLES

- Sensor As an element that senses a variation in input energy to produce a variation in another or same form of energy is called a sensor.
- Transducer uses transduction principle to convert a specified measurand into usable output.
- Thus, a properly cut piezoelectric crystal can be called a sensor whereas it becomes a transducer with appropriate electrodes and input/output mechanisms attached to it.
- The principles can be grouped according to the form of energy in which the signals are received and generated.

Classification of transducers

* Types of energy form



ENERGY TYPES AND CORRESPONDING MEASURANDS

- Mechanical :Length, area, volume, force, pressure, acceleration, torque, mass flow, acoustic intensity, and so on ..
- Thermal :Temperature, heat flow, entropy, state of matter.
- Electrical :Charge, current, voltage, resistance, inductance, capacitance, dielectric constant, polarization, frequency, electric field, dipole moment, and so on.
- Magnetic :Field intensity, flux density, permeability, magnetic moment, and so forth.
- Radiant :Intensity, phase, refractive index, reflectance, transmittance, absorbance, wavelength, polarization, and so on.
- Chemical :Concentration, composition, oxidation/reduction potential, reaction rate, pH, and the like.

CLASSIFICATION

- It is very difficult to classify sensors under one criterion and hence, different criteria may be adopted for the purpose.

Some of these include:

1. **Transduction principles** using physical or chemical effects,
2. **Primary input quantity -** Measurand,
3. **Material and technology**,
4. **Application**
5. **Property**.

Subclassification

- Another very preliminary classification or subclassification as it may be, is based on the energy or power supply requirements of the sensors.
- This means that some sensors require power supply and there are some that do not.
- As a result, the sensors are called active and passive respectively. Sometimes, terms such as self-generating and modulating are used to qualify these.

TABLE 1. 1 ENLISTS PHYSICAL AND CHEMICAL TRANSDUCTION PRINCIPLES ALONG WITH SOME ELABORATIONS.

Table 1.1 Physical and chemical transduction principles

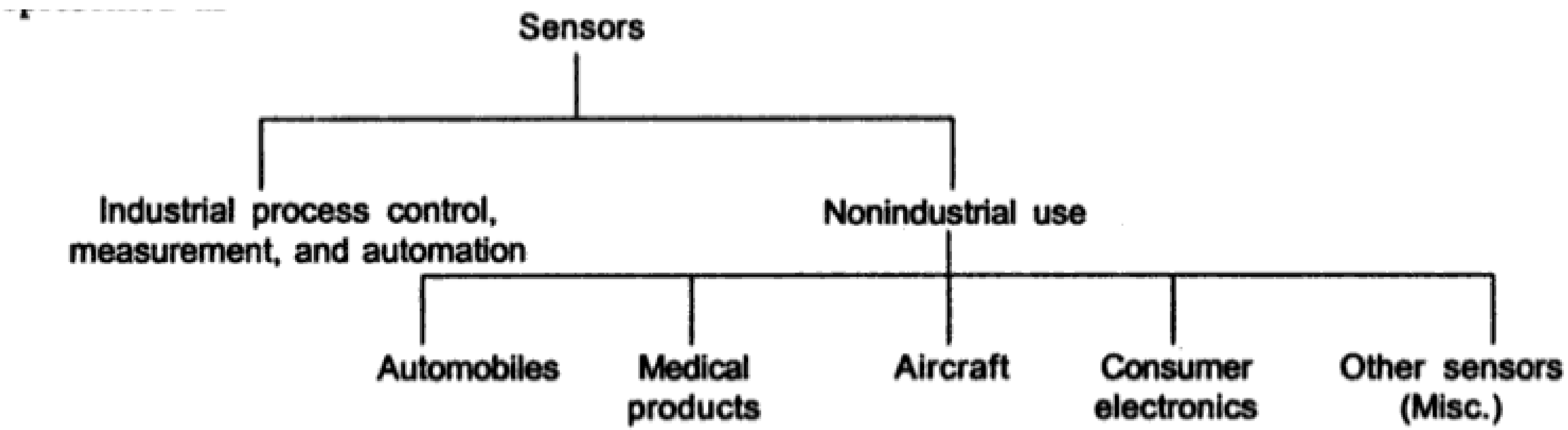
Input \ Output	*Mechanical*	*Thermal*	*Electrical*	*Magnetic*	*Radiant*	*Chemical*
Mechanical	Mechanical including acoustic effects. eg: diaphragm.	Friction effects, cooling effects. eg: thermal flowmeter.	Piezoelectricity, piezoresistivity, resistive, inductive, and capacitive changes.	Piezomagnetic effects.	Photoelasticity, interferometry, Doppler effect.	—
Thermal	Thermal expansion. eg: expansion thermometry.	—	Seebeck effect, pyroelectricity, thermoresistance. eg: Johnson noise.	—	Thermo-optical effects. eg: liquid crystals, thermo-radiant emission.	Thermal dissociation, thermally induced reaction.
Electrical	Electrokinetic effects. eg: inverse piezoelectricity.	Peltier effect, Joule heating.	Charge controlled devices, Langmuir probe.	Biot-Savart's electromagnetic law.	Electroluminescence, Kerr effect.	Electrolysis, electrically induced reaction. eg: electromigration.
Magnetic	Magnetostriction, magnetometers.	Magnetothermal effects (Righi-Leduc effect).	Ettinghaussen-Nernst effect, Galvanomagnetic effect. eg: Hall effect, magnetoresistance.	—	Magneto-optical effects. eg: Faraday effect, Cotton-Mouton effect.	—
Radiant	Radiation pressure.	Bolometer, thermopile.	Photoelectric effects. eg: photovoltaic cell, LDR's.	—	Photorefractivity, photon induced light emission.	Photodissociation, photosynthesis.
Chemical	Photoacoustic effect, hygrometry.	Thermal conductivity cell, calorimetry.	Conductimetry, potentiometry, voltametry, flame-ionization, chem FET.	Nuclear magnetic resonance.	Spectroscopy. eg: emission and absorption types, Chemiluminescence.	—

TECHNOLOGY BASED CLASSIFICATION

- Conventional sensors are now aptly supported by technologies which have yielded Micro Electro Mechanical Sensors (MEMS), CMOS image sensors, displacement and motion detectors and biosensors.

- Similarly, Coriolis, magnetic and ultrasonic flowmeters, photoelectric, proximity, Hall effect, infrared, integrated circuit (IC), temperature, radar-based level sensors are also relatively modern.

APPLICATION BASED CLASSIFICATION



PROPERTY BASED CLASSIFICATION

Table 1.3 Property based classification

	Property				
	Flow	*Level*	*Temperature*	*Pressure*	*Proximity and displacement*
Technology	Differential pressure, positional displacement, vortex, thermal mass, electromagnetic, Coriolis, ultrasonic, anemometer, open channel.	Mechanical, magnetic, differential pressure, thermal displacement, vibrating rod, magnetostrictive, ultrasonic, radio frequency, capacitance type, microwave/radar, nuclear.	Filled-in systems, RTDs, thermistors, IC, thermocouples, inductively coupled, radiation (IR).	Elastic, liquid-based manometers, inductive/LVDT, piezoelectric, electronic, fibre optic, MEMS, vacuum.	Potentiometric, inductive/LVDT, capacitive, magnetic, photoelectric, magnetostrictive, ultrasonic.

Table 1.3 (*cont.*)

	Property				
	Acceleration	*Image*	*Gas and chemical*	*Biosensors*	*Others*
Technology	Accelerometers, gyroscopes.	CMOS, CCDs (charge coupled devices).	Chemical bead, electrochemical, thermal conductance, paramagnetic, ionization, infrared, semiconductor.	Electrochemical, light-addressable potentiometric (LAP), surface plasmon resonance (SPR), resonant mirror	Mass, force, load, humidity, moisture, viscosity.

Table 1.4 shows the emerging sensor technologies with current and future application schedules as a chart.

Table 1.4 Emerging sensor technologies

	Sensors			
	Image sensors *Technology: CMOS-based*	*Motion detectors* *Technology: IR, ultrasonic, microwave/radar*	*Biosensors* *Technology: electrochemical*	*Accelerometers* *Technology: MEMS-based*
Applications	Traffic and security surveillance, blind-spot detection as autosensors (robots etc.), video conferencing, consumer electronics, biometrics, PC imaging	Obstruction detection (robots, auto), security detection (intrusion), toilet activation, kiosks videograms and simulations, light activation	Water testing, food testing (contamination detection), medical care device, biological warfare agent detection	Vehicle dynamic system (auto), patient monitoring (including pace makers etc.)

SESSION 2

CHARACTERISTICS OF MEASUREMENT SYSTEMS

- Static characteristics
  1. Accuracy
  2. Precision
  3. Resolution and
  4. Sensitivity

1.4.1 Characteristics

Sensors like measurement systems have two general characteristics, namely 1. static, and 2. dynamic.

Static characteristics

(a) *Accuracy specified by inaccuracy* or *usually error:* which is given by

$$\varepsilon_a\% = \frac{x_m - x_t}{x_t} \times 100 \tag{1.1}$$

where

t stands for true value,
m for measured value, and
x stands for measurand.

This is often expressed for the full scale output (fso) and is given by

$$\varepsilon_{fso}\% = \frac{x_m - x_t}{x_{fso}} \times 100 \tag{1.2}$$

Obviously,

$$|\varepsilon_{fso}| \le |\varepsilon_a|$$

For multi-error systems, the overall performance in terms of error can be assessed either through (i) the worst case approach which assumes that all errors add up in the same direction so that the overall error is very high, being the linear sum of all the performance errors, or through (ii) the root mean square approach which is optimistic as well as practical, when the total performance error is assessed as

$$\varepsilon_0 = \left[\sum_i (\varepsilon_i)^2\right]^{1/2} \tag{1.3}$$

(b) *Precision:* describes how far a measured quantity is reproducible as also how close it is to the true value.

The term 'repeatability' is close to precision which is the difference in output y at a given value of the input x when obtained in two consecutive measurements. It may be expressed as % FSO. Figure 1.1 shows the plot of repeatability.

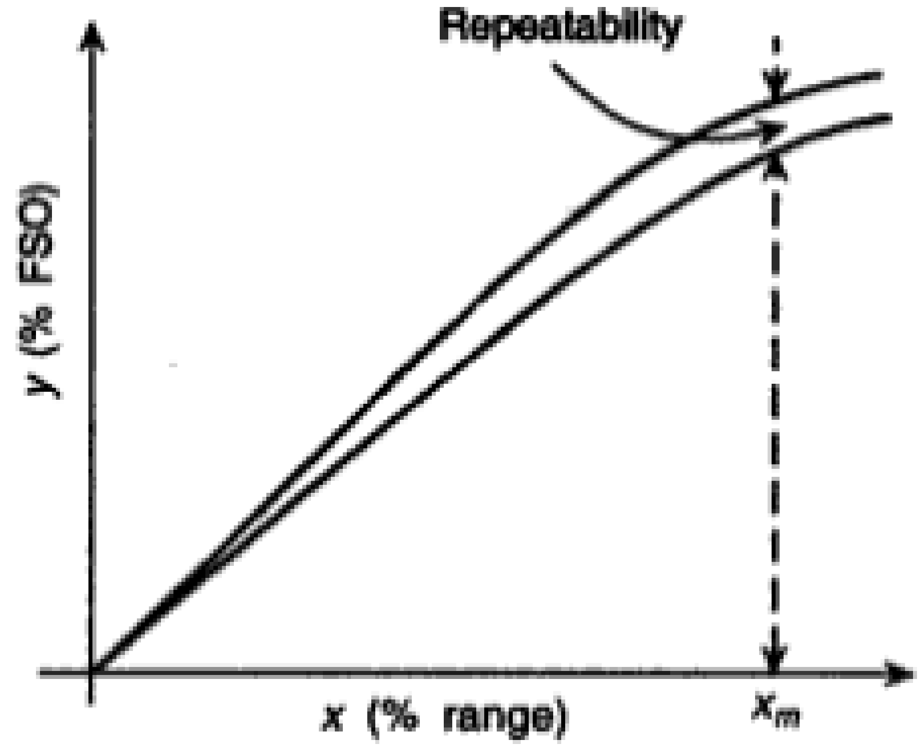


Fig. 1.1 Repeatability in y–x coordinates.

(c) *Resolution:* is defined as the smallest incremental change in the input that would produce a detectable change in the output. This is often expressed as percentage of the measured range, *MR*. The measured range is defined as the difference of the maximum input and the minimum input, that is, $MR = x_{max} - x_{min}$. For a detectable output Δy, if the minimum change in x is $(\Delta x)_{min}$, then the maximum resolution is

$$R_{max}(\%) = \frac{100(\Delta x)_{min}}{MR} \tag{1.4}$$

Over the range of operation, an average resolution has also been defined as

$$R_{av}(\%) = 100\frac{\sum_{1}^{n} \Delta x_i}{n \cdot MR} \tag{1.5}$$

(d) *Minimum Detectable Signal (MDS):* Noise in a sensor occurs because of many reasons—internal sources or fluctuations due to externally generated mechanical and electromagnetic influences. Noise is considered in detail, on individual merits and often an equivalent noise source is considered for test purposes.

If the input does not contain any noise, the minimum signal level that produces a detectable output from the sensor is determined by its noise performance or noise characteristics. For this, the equivalent noise source is connected to the input side of the ideal noiseless sensor to yield an output which is the actual output level of the sensor. The MDS is then taken as the RMS equivalent input noise. When signal exceeds this value, it is called a detectable signal.

(e) *Threshold:* At the zero value condition of the measurand, the smallest input change that produces a detectable output is called the threshold.

(f) *Sensitivity:* It is the ratio of the incremental output to incremental input, that is

$$S = \frac{\Delta y}{\Delta x} \tag{1.6}$$

In normalized form, this can be written as

$$S_n = \frac{\Delta y/\Delta x}{y/x} \tag{1.7}$$

If sensitivity or the output level changes with time, temperature and/or any other parameters without any change in input level, drift is said to occur in the system which often leads to instability.

(g) *Selectivity and specificity:* The output of a sensor may change when afflicted by environmental parameters or other variables and this may appear as an unwanted signal. The sensor is then said to be *non-selective*. It is customary to define selectivity or specificity by considering a system of n sensors each with output $y_k (k = 1, 2, \ldots, n)$. The partial sensitivity S_{jk} is defined as the measure of sensitivity of the kth sensor to these other interfering quantities or variables x_j as

$$S_{jk} = \frac{\Delta y_k}{\Delta x_j} \tag{1.8}$$

A selectivity matrix would thus be obtained with S_{jk} as the jkth entry. Obviously, an ideally selective system will have only diagonal entries S_{jj} in the selectivity matrix. An ideally specific system is characterized by having a matrix with a single entry in the diagonal. Following relationship describes selectivity, λ;

$$\lambda = \min\left[\frac{S_{jj}}{\sum_{k=1}^{n}|S_{jk}| - |S_{jj}|}\right] \qquad j = 1, 2, \ldots, n \tag{1.9}$$

Thus, for a selective group, denominator tends to zero and $\lambda \to \infty$. Also, specificity is a special case of selectivity.

(h) *Nonlinearity:* Deviation from linearity, which itself is defined in terms of superposition principles, is expressed as a percentage of the full scale output at a given value of the input. Nonlinearity can, however, be specified in two different ways, namely (i) deviation from best fit straight line obtained by regression analysis, and (ii) deviation from a straight line joining the end points of the scale. These are shown in Figs. 1.2(a) and (b). The maximum nonlinearity in the first method is always less than the maximum nonlinearity in the second one. The figure is actually half.

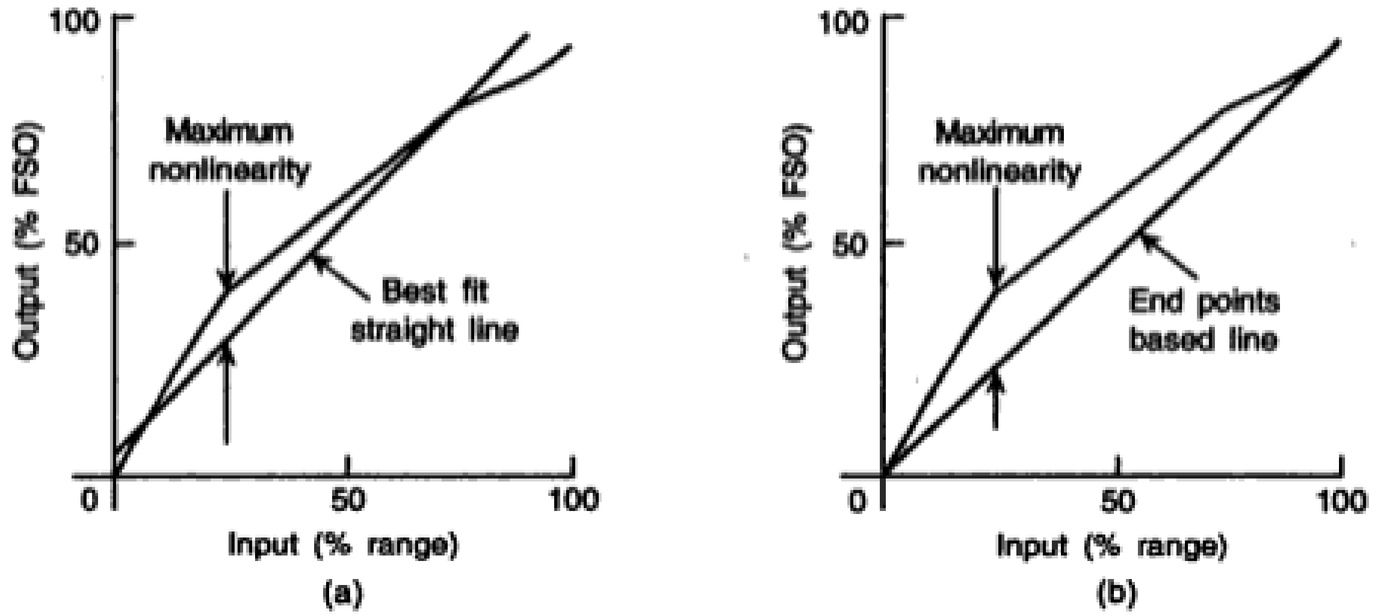


Fig. 1.2 Nonlinearity with (a) best-fit characteristics and (b) terminal-based characteristics.

A consequence of nonlinearity is *distortion* which is defined as the deviation from an expected output of the sensor or transducer. It also occurs due to presence of additional input components. If deviation at each point of the experimental curve is negligibly small from the corresponding point in the theoretical curve or from a curve made by using least square or other standard fits, the sensor is said to have *conformance* which is quantitatively expressed in % FSO at any given value of the input.

(i) *Hysteresis:* It is the difference in the output y of the sensor for a given input x when x reaches this value in upscale and downscale directions as shown in Fig. 1.3. The causes

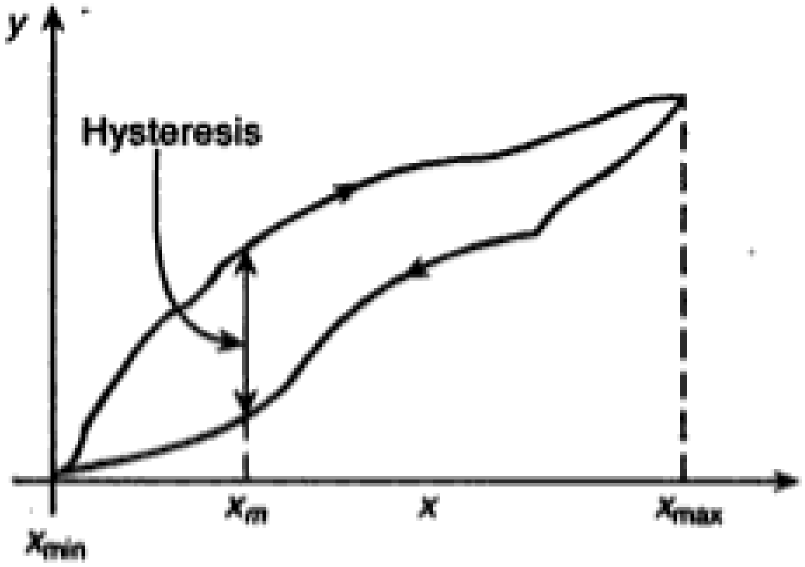


Fig. 1.3 The hysteresis curve.

It can be defined as the **maximum difference in the output values obtained by covering the measurand range** first in the increasing direction (i.e. from zero to 100%) and then in the decreasing direction (i.e. from 100% to zero).

are different for different types of sensors. In magnetic types, for example, it is the lag in alignment of the dipoles, in semiconductor types it is the injection type slow traps producing the effect, and so on.

SESSION 3

- CHARACTERISTICS OF MEASUREMENT SYSTEMS
  - Dynamic characteristics

- ENVIRONMENTAL PARAMETERS

Dynamic caracteristics

These involve determination of transfer function, frequency response, impulse response as also step response and then evaluation of the time-dependent outputs. The two important parameters in this connection are (a) fidelity determined by dynamic error and (b) speed of response determined by lag.

For determining the dynamic characteristics, different specified inputs are given to the sensor and the response characteristics are studied. With step input, the specifications in terms of the time constant of the sensor are made. Generally, the sensor is a single time constant device and if this time constant is τ, then one has the specifications as given in Table 1.5.

Table 1.5 % Response time of the sensors

% Response time	*Value in terms of τ*
$t_{0.1}$ or 10	0.104τ
$t_{0.5}$ or 50	0.693τ
$t_{0.9}$ or 90	2.303τ

This gives $t_{0.9}/t_{0.5} = 3.32$ which is taken as a quick check relation.

Impulse response as well as its Fourier transform are also considered for time domain as well as frequency domain studies.

1.5 ENVIRONMENTAL PARAMETERS (EP)

These are the external variables such as temperature, pressure, humidity, vibration, and the like which affect the performance of the sensor. These parameters are not the ones that are to be sensed.

For non-temperature transducers, temperature is the most important environmental parameter (EP). For any EP, the performance of the transducer can be studied in terms of its effect on the static and dynamic characteristics of the sensor as has already been discussed. For this study, one EP at a time is considered variable while others are held constant.

SESSION 4, 5, 6

SESSION 4: Characterization and its type - Electrical characterization,

SESSION 5: Mechanical Characterization and Thermal Characterization

SESSION 6: 1. Optical Characterization.

2. Errors and its classification,

CHARACTERIZATION

Characterization of the sensors **can be** done in many ways depending on the types of sensors. specifically microsensors These are

Electrical,

Mechanical,

Optical,

Thermal,

Chemical,

Biological and so on.

ELECTRICAL CHARACTERIZATION

It consists of evaluation of electrical parameters

1. IMPEDANCES, VOLTAGE AND CURRENTS

- The Knowledge of sensor "output impedance" is important for coupling measuring instruments.
- For voltage sensitive sensors, the ratio of the input impedance of the measuring equipment to the output of the transducer/sensor should be very high, for current sensitive sensors. reverse is true.

2. BREAKDOWN VOLTAGES AND FIELDS

'Breakdown' of the insulating parts of the sensor is very critical as the health of the system depends on it.

For Metal – Insulator – Metal (MIM) or for Metal – Insulator -Semiconductor (MIS) Structures, the breakdown of the insulating film is studied by the system.

of Fig. 4

Three different types of breakdown are

i) Dielectric strength.

ii) wear out and

iii) Current induced breakdown.

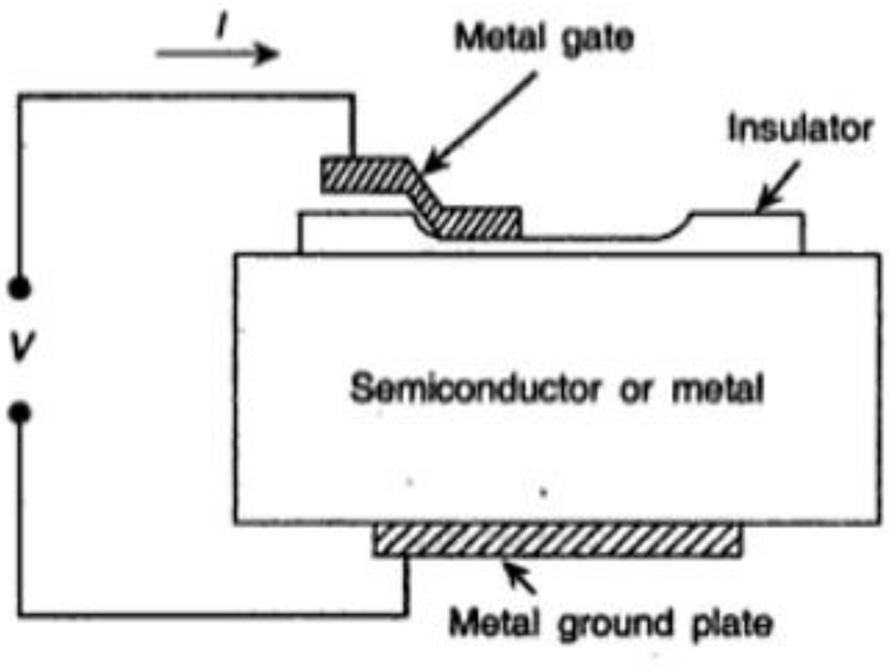


Fig. 1.4 Structure of a metal oxide semiconductor.

- Breakdown generally implies a sudden or 'avalanche' change in the voltage or current dropping to a negligible value and current rising to a very high value.
- Breakdown may be extrinsic or intrinsic though the mechanism in either case is basically the same.
- There occurs a high local field in the material which may be defect-induced which then is called *extrinsic*.
- However. if this is high field induced microvoids is called *intrinsic* type.

In the latter *case*, the high field induces microvoids to generate defects leading it to behave as the extrinsic type.

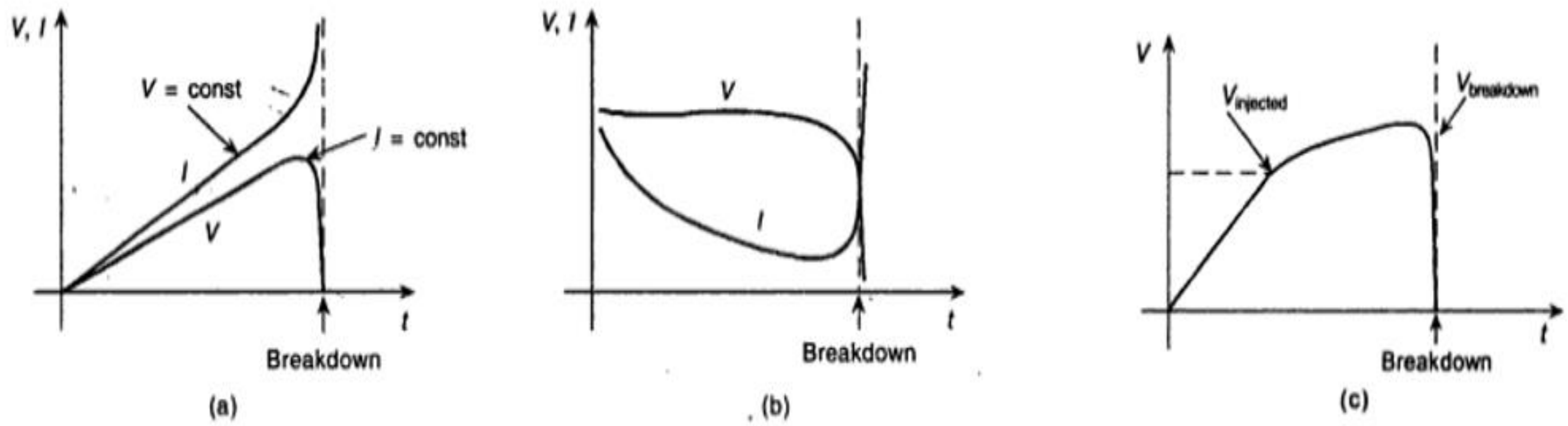


Fig. 1.5 Breakdown characteristics: (a) dielectric strength, (b) wear out, (c) current induced type.

3. LEAKAGE CURRENT

Leakage current measurement specifies the sensor quality, specifically its insulating quality as also the quality of p-n junctions wherever it exists.

4. NOISE

- It comes from electromagnetic interference. AC magnetic fluctuation, 50Hz supply pick up. mechanical or acoustical vibration, or photon induced output.
- Sensors are to be characterized for noise testing for immunity to such noise. For testing purposes, different noise sources are developed.

5. CROSSTALK

- In multichannel or array sensors "crosstalk" may occur due be overlapping of signals between the two adjacent transducer elements.
- It may however occur in a single transducer system because of inductive or capacitive coupling or coupling through the common voltage source during transduction inside the element. It is done using correlation techniques.

MECHANICAL AND THERMAL CHARACTERISATION

It involves mechanical and thermal properties related to the overall reliability and integrity of the transducer. as well as relevant transduction process.

- **Reliability is an important aspect of characterization. By means of testing, the functional and reliable portion of a batch of sensors or transducers is identified.**
- **Basically, failure analysis is performed and the mechanism of failure is attempted to be eliminated and thereby reduce the subsequent failures.**

In fact, the above two approaches are supplementary to each other.

Failure of transducers can be divided into three different categories.

(i) Catastrophic early life failures. often called infant mortality.

(ii) Short term drifts in sensor parameters and

(iii) Long term drifts and failures.

Catastrophic failure of the sensors is the complete failure in the normal operation. It is called wear out if it occurs in later life.

Short term and Long term drifts are, in effect, changes in sensor parameters and are, therefore to be studied more intensely for the sensor characterization.

The Reliability of an item is given by what is known as reliability function *R(x)* which is the probability that the item would survive *for* stated interval, say between 0 and *x*.

If F(x) is the probability of failure. then

$$R(x) = 1 - F(x) = \int_{x}^{+\infty} f(t)\,dt$$

$$= \frac{\text{No. of 'sound' components at instant } x}{\text{Total no. of components at } x = 0}$$

The probability of failure F(x) is actually a cumulative distribution function and in reliability statistics, the distribution functions that are used may have the following characteristics.

(i) Normal

(ii) Exponential

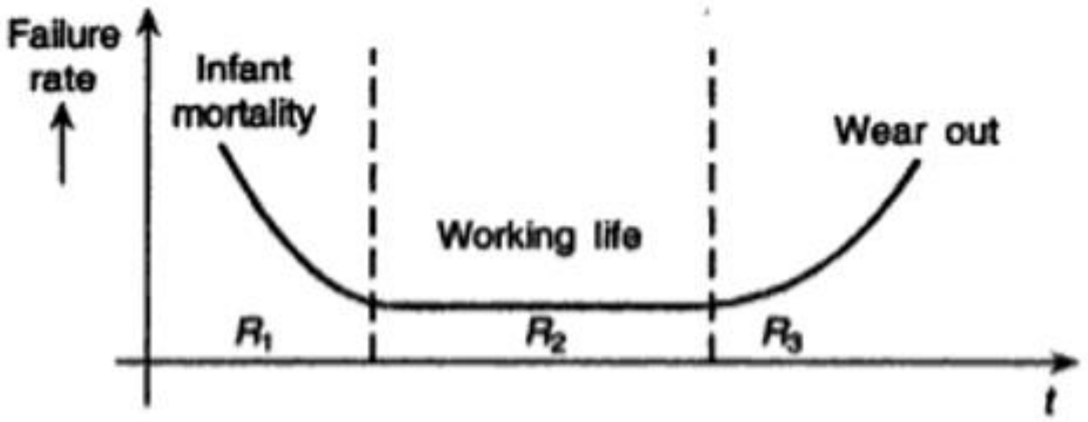
(iii) Log – Normal

(iv) Gamma and

(V) Weibull.

In general, the failure has a time dependency as shown in Fig 1.6. The curve between failure rate and time appears like a bath tub which can be divided into three distinct zones

- Zone 1 is the infant mortality zone.
- *Zone* 2 is the working zone with constant failure rate.
- Zone 3 is the wear out zone.



- **As the normal working life is usually very long, Estimating failure is tested by 'Accelerated ageing test'.**
- **Before this ageing test, screening steps are taken to isolate the defective transducers. These steps *vary depending* on the types of the transducers. For Standard Integrated Type Sensors (SITS), the typical tests that are performed are briefly discussed.**

***a. High Temperature Burn In*:**

The sensors are subjected to a high temperature over a stipulated period, Usually 125 degree Celsius for 48 hours for SITS, when the defective units are burnt out and the remaining ones are expected to run for the expected life.

b. High temperature storage bake:

The units are baked at a high temperature, usually at 250 degree Celsius for SITS, for *several hours*, when the Instability mechanisms such as contamination, bulk defects, and metallization problems are enhanced in some units which were initially defective. These units are then screened out

c. Electrical overstress test:

Where progressively larger voltages up to 50% in excess of specification are applied over different intervals of time so that failures due to insulation. interconnection or oxide formation can occur in some units which were originally defective and screened out.

d. Thermal shock test:

Mainly done for packaging defects, where the units are subjected to a temperature between 65 degree and 125 degree Celsius for about 10 seconds for every temperature. The time is gradually increased to 10 minutes and the cycle is repeated 10 times. The failed units are rejected.

e. Mechanical shock test:

Also for packaging, this lest is performed by dropping the units from a specified height that varies from 3 to 10 m. Alternately. the unit as shaken by attaching it to a shaking table for a specified period of time.

As has already been mentioned,

Real time operational test for reliability is difficult to perform so that accelerated ageing *test* has been proposed. The test should simulate the real ageing peace in a much shorter time. High stress is imposed on the sensors and results from such test are used to predict the performance in the normally stressed condition.

The results should be interpreted for

(i) True accelerated ageing.

(ii) Valid extrapolation to obtain expected performance under normal conditions.

(iii) Determining the acceleration factor for the scaling.

that is how many hours of normal operation correspond to 1 hour of accelerated operation.

Appropriate models have been developed for the purpose and failures with respect to specified parameters such as leakage current. Temperature, and so forth are predicted.

OPTICAL CHARACTERIZATION:

It a usually done by ascertaining absorption coefficient, refractive index, reflectivity and the like. Here, again the consideration of the individual merit comes in.

ERRORS AND ITS CLASSIFICATION

No Measurement is Accurate!

- Errors occur because of:
  1. Parallax error (incorrectly sighting the measurement).
  2. Calibration error (if the scale is not accurately drawn).
  3. Zero error (if the device doesn’t have a zero or isn’t correctly set to zero).
  4. Damage (if the device is damaged or faulty).
  5. Limit of reading of the measurement device (the measurement can only be as accurate as the smallest unit of measurement of the device).

Types of Errors

Basically Three types of errors are studied:-

1. Gross Errors
2. Systematic Errors
3. Random Errors

Gross Errors

Gross Errors mainly covers the human mistakes in reading instruments and recording and calculating measurement results.

Example:- Due to oversight, The read of Temperature as 31.5˚.while the actual reading may be 21.5˚.

Gross Errors may be of any amount and then their mathematical analysis is impossible. Then these are avoided by adopting two means:-

1.Great care is must in reading and recording the data.

2.Two , Three or even more reading should be taken for the quantity under measurement.

Systematic Errors

Systematic Errors classified into three categories :-

1. Instrumental Errors
2. Environmental Errors
3. Observational Errors

Instrumental Errors

These errors arises due to three main reasons.

1. Due to inherent shortcoming in the instrument.

   Example:- If the spring used in permanent magnet instrument has become weak then instrument will always read high. Errors may caused because of friction , hysteresis , or even gear backlash.

2. Due to misuse of the instruments.
3. Due to Loading effects of instruments.

Environmental Errors

- These errors are due to conditions external to the measuring Device including conditions in the are surrounding the instrument.
- These may be effects of Temperature, Pressure, Humidity, Dust, Vibrations or of external magnetic or electrostatic fields.

Observational Errors

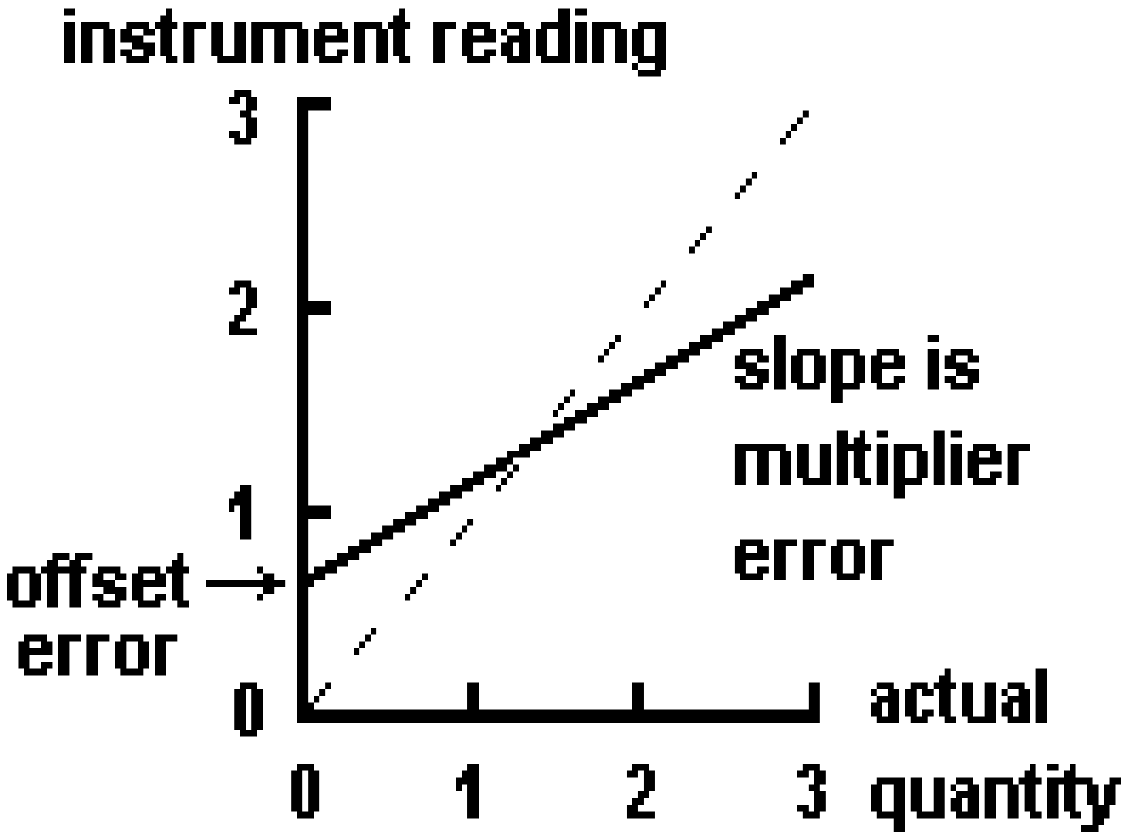
There are many sources of observational errors:-

-- Parallax, i.e. Apparent displacement when the line of vision is not normal to the scale.

-- Inaccurate estimate of average reading.

-- Wrong scale reading and wrong recording the data.

-- Incorrect conversion of units between consecutive reading.



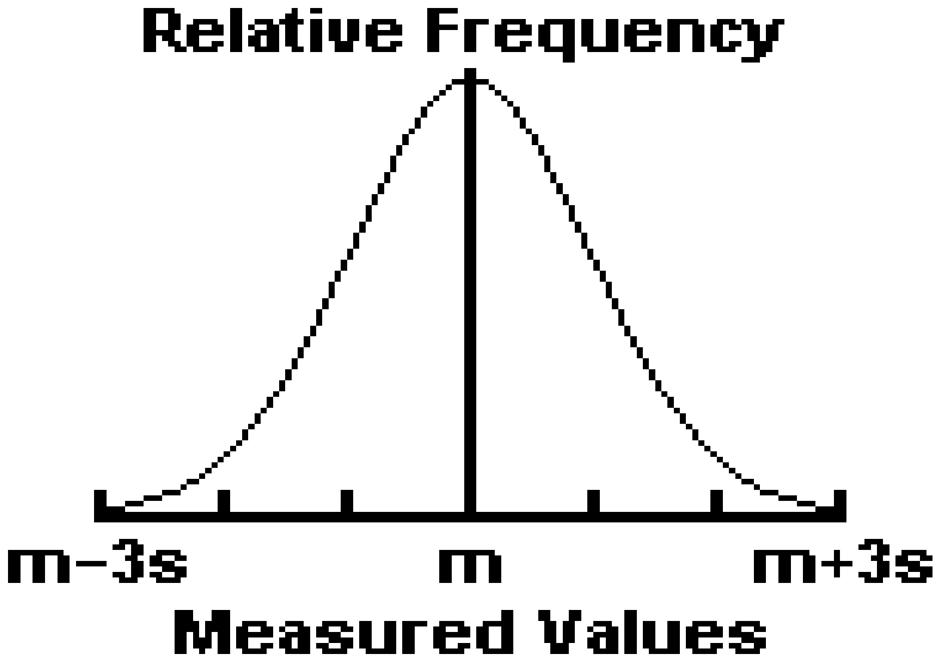
Here in Graph, Full Line shows the systematic Error in non Linear Instrument.

While Broken Line shows response of an ideal instrument without Error.

Random Errors

The quantity being measured is affected by many happenings in the universe. The errors caused by happening or disturbances about which we are unaware are Random Errors. Its also known as residual Errors.

Example of Random Error



The Gaussian normal distribution. m = mean of measurements. s = standard deviation of measurements. 68% of the measurements lie in the interval $m - s < x < m + s$; 95% lie within $m - 2s < x < m + 2s$; and 99.7% lie within $m - 3s < x < m + 3s$.

SESSION 7

- SELECTION OF TRANSDUCERS
- INTRODUCTION TO MECHANICAL SENSORS

SELECTION OF TRANSDUCERS

- There are many ways for measurement of a physical quantity. While selection of transducers, how to select a transducer for a particular application, following points should be kept in the mind.

- Unfortunately most transducers are not sensitive to just one quantity.

- If measurements are to be made under conditions where there is likelihood of two or more input quantities influencing the transducer, it is desirable to **select a transducer which is sensitive to the desirable quantity and insensitive to the unwanted quantity.**
- If this is not possible, ways and means should be found to eliminate or compensate for the effects of the unwanted input quantity.

SELECTION OF TRANSDUCERS

- Operating Principle.
- Sensitivity.
- Operating Range.
- Accuracy
- Cross sensitivity
- Transient and Frequency Response.
- Loading Effects.
- Environmental Compatibility
- Usage and Ruggedness
- Electrical aspects.
- Stability and Reliability

SELECTION OF TRANSDUCERS

- Operating Principle
  - The transducers are many times selected on the basis of operating principle used by them. The operating principles used may be resistive, inductive, capacitive, optoelectronic, piezoelectric etc.
- Sensitivity.
  - The transducer must be sensitive enough to produce detectable output.
- Operating Range.
  - The transducer should maintain the range requirements and have a good resolution over its entire range. The rating of the transducer should be sufficient so that it does not breakdown while working in its specified operating range.
- Accuracy
  - High degree of accuracy is assured if the transducer does not require frequent calibration and has a small value for repeatability. It may be emphasized that in most industrial applications, repeatability is of considerably more importance than absolute accuracy.

SELECTION OF TRANSDUCERS

- Cross sensitivity
  - Cross sensitivity is a further factor to be taken into account when measuring mechanical quantities.
  - There are situations where the actual quantity is being measured is in one plane and the transducer is subjected to variations in another plane.
  - More than one promising transducer design has had to be abandoned because the sensitivity to variations of the measured quantity in a plane perpendicular to the required plane has been such as to give completely erroneous results when the transducer has been used in practice.
- Transient and Frequency Response
  - The transducer should meet the desired time domain specifications like peak overshoot, rise time, settling time and small dynamic error.
  - It should ideally have a flat frequency response curve. In practice, however, there will be cutoff frequencies and higher cut off frequency should be high in order to have a wide bandwidth.

SELECTION OF TRANSDUCERS

- Loading Effects.
  - The transducer should have a high input impedance and a low output impedance to avoid loading effects.
- Environmental Compatibility
  - It should be assured that the transducer selected to work under specified environmental conditions maintains its input-output relationship and does not break down.
  - For example, the transducer should remain operable under its temperature range.
  - It should be able to work in corrosive environments (if the application so requires), should be able to withstand pressures and shocks and other interactions to which it is subjected to.
  - Insensitivity to Unwanted Signals in the sense the transducer should be minimally sensitive to unwanted signals and highly sensitive to desired signals.
- Usage and Ruggedness
  - The ruggedness both of mechanical and electrical intensities of transducer versus its size and weight must be considered while selecting a suitable transducer.

SELECTION OF TRANSDUCERS

- ❖Electrical aspects.
  - ❖The electrical aspects that need consideration while selecting a transducer include the length and type of cable required.
  - ❖Attention also must be paid to signal to noise ratio in case the transducer is to be used in conjunction with amplifiers.
  - ❖Frequency response limitations must also be taken into account.
- ❖Stability and Reliability
  - ❖The transducer should exhibit a high degree of stability to be operative during its operation and storage life.
- ❖In general, the transducer should maintain the expected input output relationship as described by its transfer function so as to avoid errors in transducers.

MECHANICAL SENSOR

- Mechanical sensors, are those which have a mechanical quantity as the input and the output may be a quantity such as an electrical, magnetic, optical, thermal, and so on. In such a case, motion, displacement, speed, velocity, force, acceleration, torque and other such quantities should be measured by mechanical sensors.
- Process variables like pressure, flow, and level should also be considered as mechanical inputs and sensors for measurement of such variables should also be considered as mechanical sensors.
- In many such sensors 'electromechanical coupling' is involved. As such, the primary objective is to convert the input form into an electrical output form for convenience of processing and display.
- In this respect, mechanical sensors are also termed as electromechanical or mechano- electrical sensors.
- It would, however, be seen that many sensors may be categorized under more than one category without being inappropriate. In the proceedings, appropriate references may be made for such multiplacement.

- Further, same sensor is often used for measuring different variables by appropriate adaptation. In this way, it is not always possible to uniquely identify a sensor for a specific function.
- ' Many mechanical variables are secondary in nature such as the ‘motion’ of the tip of a bimetallic element (thermal sensor) which is the result of the temperature variation of the element. Temperature, here, is the primary variable or input. The tip motion is angular or rotational.
- Such rotational or translational displacements are measurable by various means. But direct measurement by using a pointer attached to the ‘moving’ end often leads to poor accuracy because of small movement and/or low resolution.
- Instead, resistive potentiometers, LVDT’s, capacitive sensors, and so forth are used, not only for displacements alone but also for various other related variables.

SESSION 8

- RESISTIVE POTENTIOMETER AND TYPES,
- STRAIN GAUGE: THEORY, TYPE, DESIGN, CONSIDERATION, SENSITIVITY.

RESISTIVE POTENTIOMETERS

- Resistive potentiometer is a kind of variable resistance transducer.
- Potentiometer is essentially a voltage divider used for measuring electric potential (voltage).
- Others in this category are

➢ Strain gauges

➢ RTD

➢ Thermistor

➢ Wire anemometer

➢ Piezo resistor and many more.

Precision wire-wound potentiometer

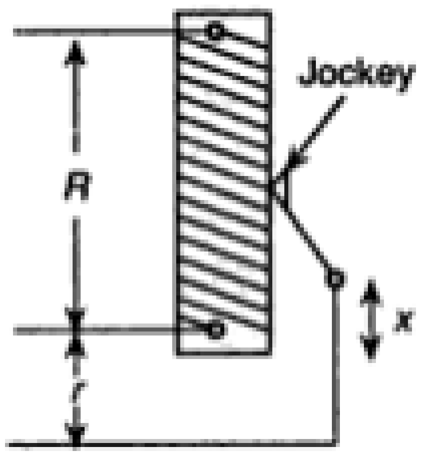
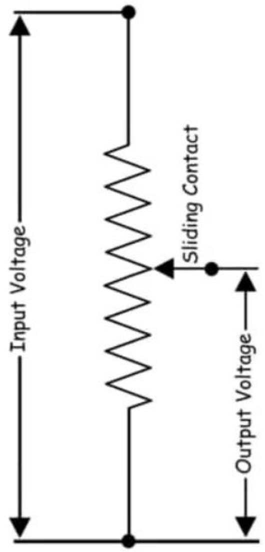
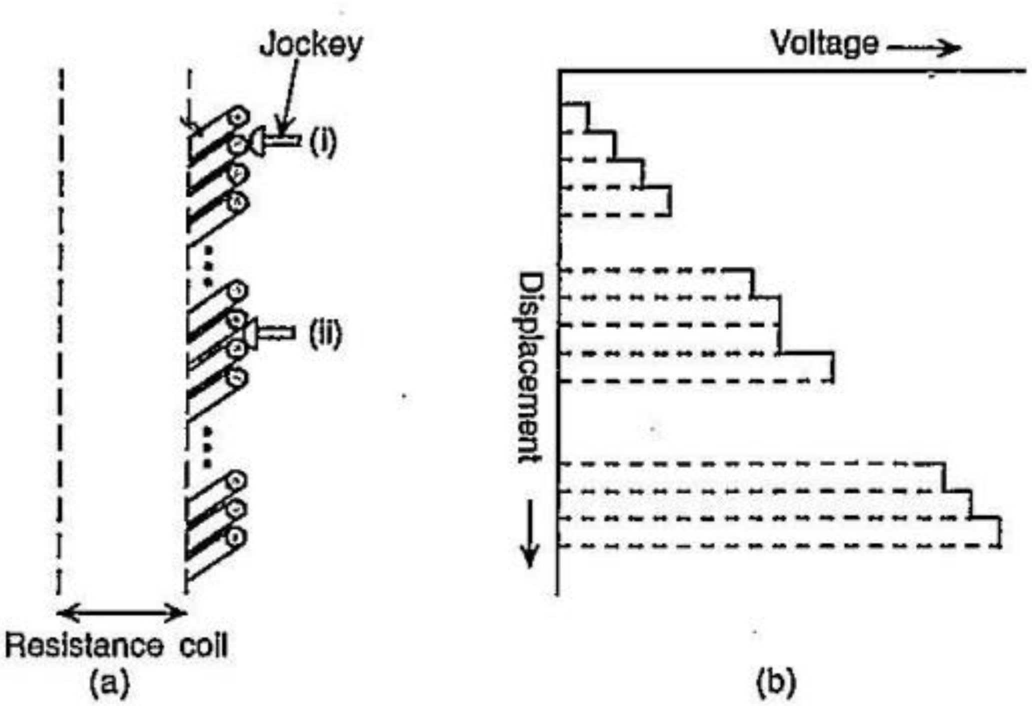


Fig. 2.1 Wire-wound potentiometer.



- Precision wire-wound potentiometer which is used as a sensor.
- A major advantage with this type is its large output.
- Resolution and noise are important aspects to be discussed in connection with it

CROSS-SECTION OF THE N-TURN WINDING



ig. 2.2 (a) Jockey contact schemes: (i) single wire contact, (ii) two-wire shorting, (b) corresponding voltage levels.

CROSS-SECTION OF THE N-TURN WINDING

2.2(a) with the wiper in two different possible positions:

(i) Touching only one wire

(ii)Touching two turns is important as is obvious.

In case (i) for a voltage supply V to the potentiometer, the voltage resolution would be

$$\Delta V = \frac{V}{n}$$

In Fig. 2.2(b), the solid line stairs show the output voltage steps each of which is equivalent to a value Vin.

But during transit, two adjacent wires are likely to be shorted as shown in Fig. 2.2(a) and a minor resolution pulse of magnitude is obtained, where pth and (p+1)th wires are shorted.

$$\Delta V_m = V_p\left[\frac{1}{n-1} - \frac{1}{n}\right]$$

This shows that with increasing value of p, minor pulse magnitude also increases and the loss in resolution due to this shorting leads to an actual resolution value.

$$\Delta V - \Delta V_m = \frac{V}{n} - V_p\left[\frac{1}{n-1} - \frac{1}{n}\right]$$

- The jockey shape/profile or the ratio of jockey radius to wire radius and geometry of wire winding should be considered for reducing ΔV_m.
- If jockey radius is small, with the jockey in use for some time with pressure, the wire gets its round surface worn out to develop a flat surface and finally gets torn.
- With a large radius of wire and close winding, this effect is small, but may short more than two wires during the movement of the jockey and hence, precision of measurement is affected.
- For circular wire and circular jockey,

  It is recommended that the ratio of their radii be around 10,

  that is, rjockey/rwire approximately equal to 10.
- Also, materials of resistance wire and jockey are equally important, particularly from the, wearing point of view and 'noise'.
- For noise, among other things, the jockey construction is to be considered seriously.

A few types of the jockeys

common.

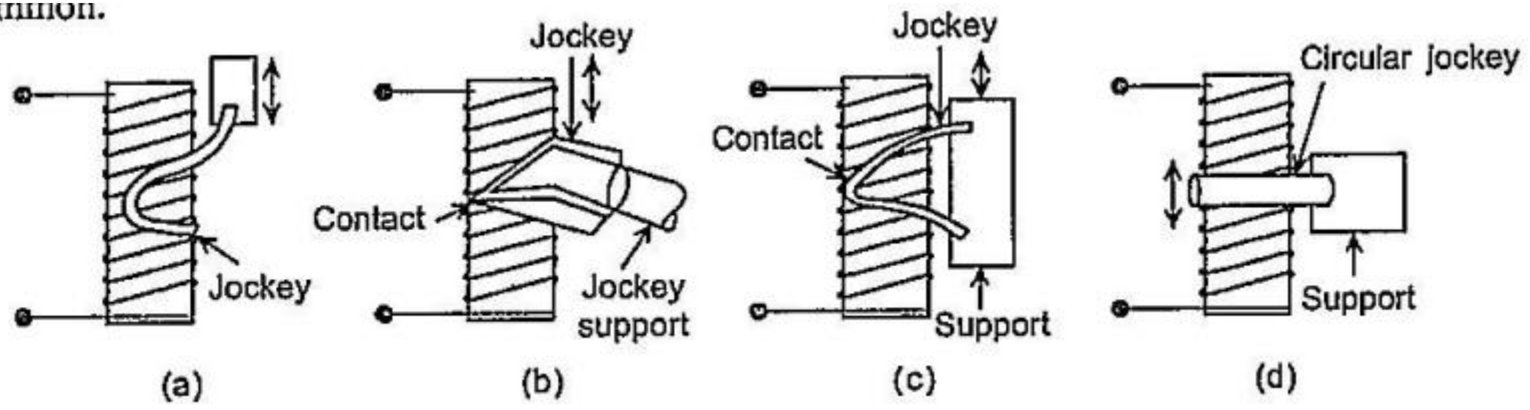


Fig. 2.3 Different designs of jockeys.

- A few types of the jockeys are shown in Figs. 2.3(a), (b), (c), and (d).
- The pressure at contact with the wire is provided by giving an adequate flexibility to the arm in relation to its mass.
- However, the required pressure is dependent on the materials, jockey to wire radius ratio, and the proposed lifespan of the potentiometer. A value of 10-50 mN is quite common.

Noise is contributed by

1. Irregularities in resolution-a random type noise,

2. Thermal motions of molecules that come in equilibrium with random motions of electrons giving rise to white/Johnson noise with equivalent voltage output as

$$\{\langle V^2\rangle\}^{1/2} = \sqrt{4kTR\Delta f}\,,$$

3.Contact non-uniformity mainly produced due to changing contact area and hence, contact resistance-aggravated by the presence of foreign particles in the area of contact (contact area changes with use, also contamination and oxidation change the resistance and hence, noise)

4.Rubbing action between the jockey and the wire-an equivalent of 100-300 µV is easily obtained with this rubbing action, and

5.Thermoelectric action specifically at high temperatures and DC operations.

Sensitivity

- Under ideal unloaded condition of the potentiometer is the output voltage per unit travel of the jockey.
- Irregularities occur

(i) at the potentiometer ends and

(ii) due to power dissipation and corresponding rise in resistance of the potentiometer.

Adequate corrections are to be made for these.

- A proper choice of the wire material with safety limit extended in current carrying capacity can minimize these errors to a certain extent.
- The performance of the potentiometer changes in the loaded condition. Specifically, linearity is badly affected.
- Considering the circuit of Fig. 2.4, if RL is the load resistance, Vi and Vo are input and output voltages respectively, R is the instantaneous tapped resistance across which Vo is obtained,
- If the jockey begins movement from the bottom end, so that minimum Ri = 0 and maximum Ri = R.

$$\frac{V_o}{V_i} = \frac{R_i/R}{1+\frac{R}{R_L}\left(1-\frac{R_i}{R}\right)}$$

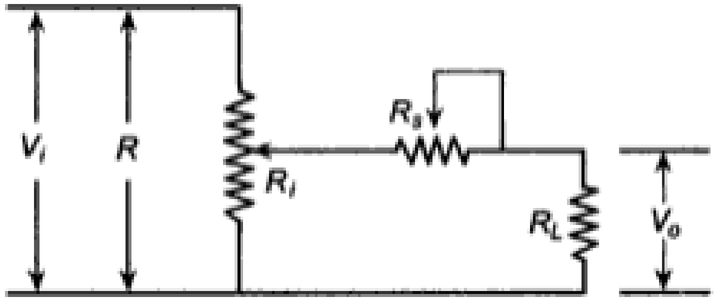


Fig. 2.4 Circuital method of drawing output from the potentiometer (for better linearity).

The figure also shows a variable series resistance R_s which is, in fact, optional and Eq. (2.4) has been obtained with $R_s = 0$. Generally, for ideal condition, $V_{oi}/V_i = R_i/R$. Representing R_i/R by ρ and R/R_L by λ, the percentage error in output–input voltage ratio is given as

$$\varepsilon = \frac{(V_{oi}/V_i - V_o/V_i)}{(V_{oi}/V_i)} \times 100$$

$$= \left(1 - \frac{1}{(1 + \lambda\rho(1-\rho))}\right) \times 100 \qquad (2.5)$$

Plots of ε versus ρ can be drawn with λ as a parameter, using Eq. (2.5), to show that the percentage deviation from linearity may be as high as 20% at $R_i = R/2$ for $R_L = R$. However, this is kept to within 1% by making $R_L \geq 20R$.

Alternate methods make use of

(i) a potentiometer which itself has nonlinear characteristics or
(ii) a nonlinear variable resistance R_s in series with the load.

The first method, in effect, proposes a design of the former, on which the winding is made, to have a nonlinear profile on the side the jockey moves. This nonlinear profile is such that the resistance ratio R_i/R curve drawn against the jockey movement (travel) is complementary to that of V_o/V_i.

In the second method, since R_s is also variable, a double jockey system—one for R and the other for R_s with equal lengths to move should be used. It can be shown that a resistance $R_s = R/4$ with parabolic resistance characteristics about an axis of symmetry at $x = 0.5$ are necessary for the purpose, where x is the normalized movement from 0 to 1.

As has already been mentioned, materials, both of the wire and the jockey are equally important. Table 2.1 shows a list of materials for the wire and the jockey which can be used in correspondence.

Table 2.1 Materials for wire and jockey

Wire	Jockey
1. Copper–nickel alloys like constantan (Cu 55–Ni 45), advance, ferry alloy, eureka and so on.	(a) Gold, gold–silver, (b) Ni 40–Ag 60, 10% graphite in Cu or 2–5% graphite in Ag.
2. Nickel–chromium alloys such as nichrome (Ni 80, Cr 20), Karma and so forth.	Group (b) above, and/or Rh or Rh-plated metals, gold–silver, osmium–iridium, Cu 40–Pd, ruthenium 10–Pt, Gold.
3. Silver–palladium alloys	Pt–iridium, Au 10–Cu 13–Ag 30–Pd 47.
4. Platinum–iridium	Pt–iridium

The wire is precision-drawn and annealed in a reducing atmosphere. The resistance per unit length varies from 0.25–1.5 μΩm. The temperature coefficient of resistance is material-dependent and lies between 2×10^{-5}/°C and 10^{-4}/°C. Wire diameter tolerance is prescribed to be less than 5% at 0.025 mm.

STRAIN GAUGE

- In early mid- nineteenth century, the basic principle of strain gauge is change in resistance of a metallic wire in response to strain produced. But now the application is so vast, that it is difficult to prepare a gist of principle.
- Strain gauges are of two types

1)Resistance type

2) Semiconductor type - more recent origin.

RESISTANCE TYPE:

- Resistance strain gauges can be divided into two categories-

(a) Unbonded

(b) Bonded

The former, being of limited use has received less attention than the latter.

Unbonded strain gauge

- Unbonded strain gauge consists of a piece of wire stretched in multiple folds between a pair or more of insulated pins fixed to movable- members of a 'body' or even a single flexible member whose strain is to be measured.
- There occurs a relative motion between the two members on strain and the wire gets strained as well with a corresponding change in its resistance value. The scheme of such a system is shown in Fig. 2.5.

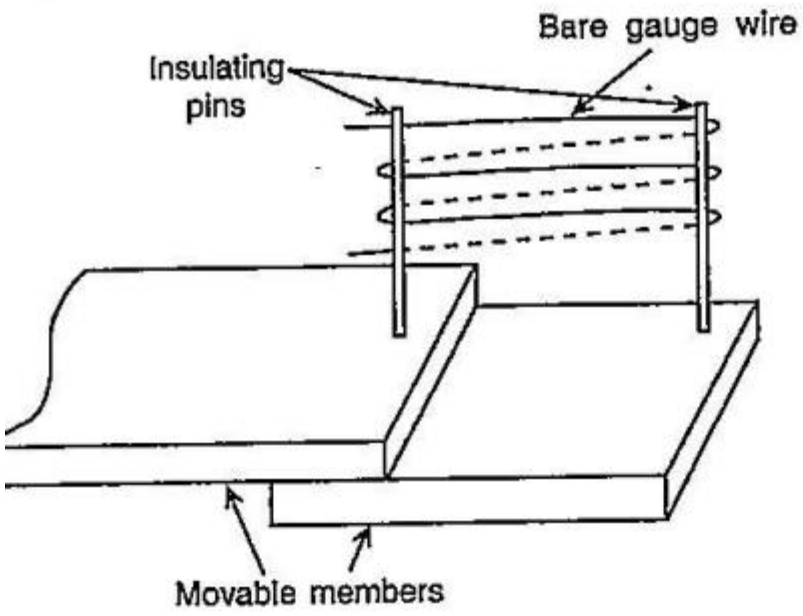


Fig. 2.5 Mounting an unbonded strain gauge.

Bonded Type

- The bonded type is more common and in its simplest form consists of wire/strip of resistance material arranged usually in the form of a grid for larger length and resistance value.
- The grid is bonded to the test specimen with an insulation layer between the gauge material and the specimen as shown in Fig. 2.6.
- If the insulation and the bonding material thickness is h which also is the height of the wire above the specimen surface and H is the distance of the neutral axis of the specimen from its surface, then the actual strain ε, in terms of measured strain ε_m, is given by

$$\varepsilon = \varepsilon_m \frac{H}{h+H}$$

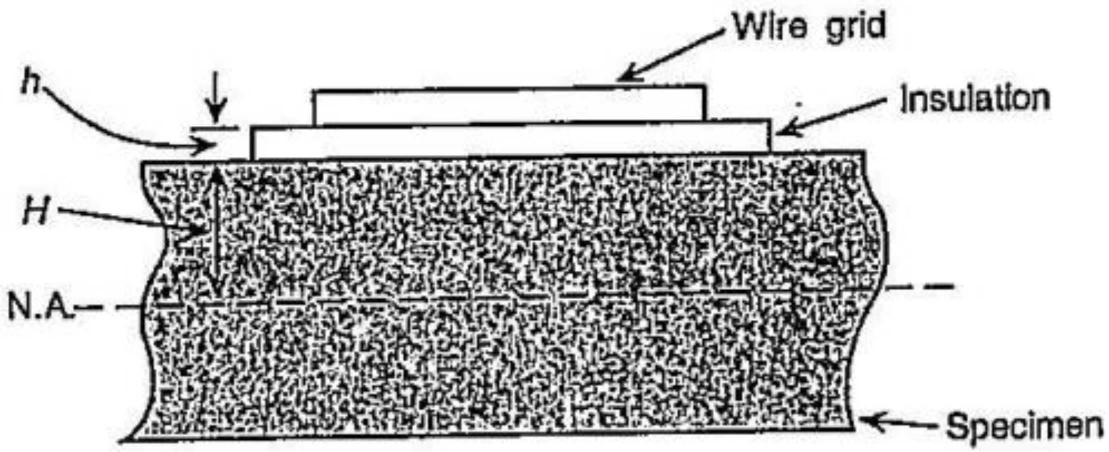


Fig. 2.6 Mounting a bonded strain gauge.

Classification of resistance gauges

Depending upon the implementation, the resistance gauges can be classified as:

(a) Unbonded metal wire

(b) Bonded metal wire

(c) Bonded metal foil

(d) Thin metal film by vacuum deposition

(e) Thin metal film by sputter deposition

- Considering a circular cross-section metal resistance wire of length l and cross-sectional area A with resistivity ϐ of the material, the unstrained resistance of the wire is given by $R = \frac{\rho l}{A}$
- If the wire is uniformly stressed along its length (Fig. 2.7) and if the stress is given by σ, then

$$\frac{dR}{d\sigma} = \frac{d}{d\sigma}\left(\rho \frac{l}{A}\right)$$

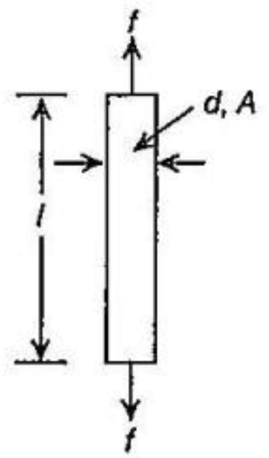


Fig. 2.7 Straining of an elastic member.

which gives

$$\left(\frac{1}{R}\right)\frac{dR}{d\sigma} = \left(\frac{1}{l}\right)\left(\frac{\partial l}{\partial \sigma}\right) - \left(\frac{1}{A}\right)\left(\frac{\partial A}{\partial \sigma}\right) + \left(\frac{1}{\rho}\right)\left(\frac{\partial \rho}{\partial \sigma}\right)$$

Eliminating all σ terms, we get

$$\frac{dR}{R} = \frac{\partial l}{l} - \frac{\partial A}{A} + \frac{\partial \rho}{\rho} \tag{2.9}$$

If the wire has a diameter d then the lateral contraction of the wire, $\Delta d/d = (1/2)\,(dA/A)$, is related to the fractional extension of the length, $\varepsilon = \Delta l/l$ by the Poisson's ratio μ as

$$\frac{\Delta d}{d} = -\frac{\mu \Delta l}{l} \tag{2.10}$$

so that Eq. (2.9) changes to

$$\frac{\Delta R}{R} = (1 + 2\mu)\frac{\Delta l}{l} + \frac{\Delta \rho}{\rho} \tag{2.11}$$

The strain sensitivity or the gauge factor λ is now defined as the ratio $(\Delta R/R)/(\Delta l/l)$ and is en by

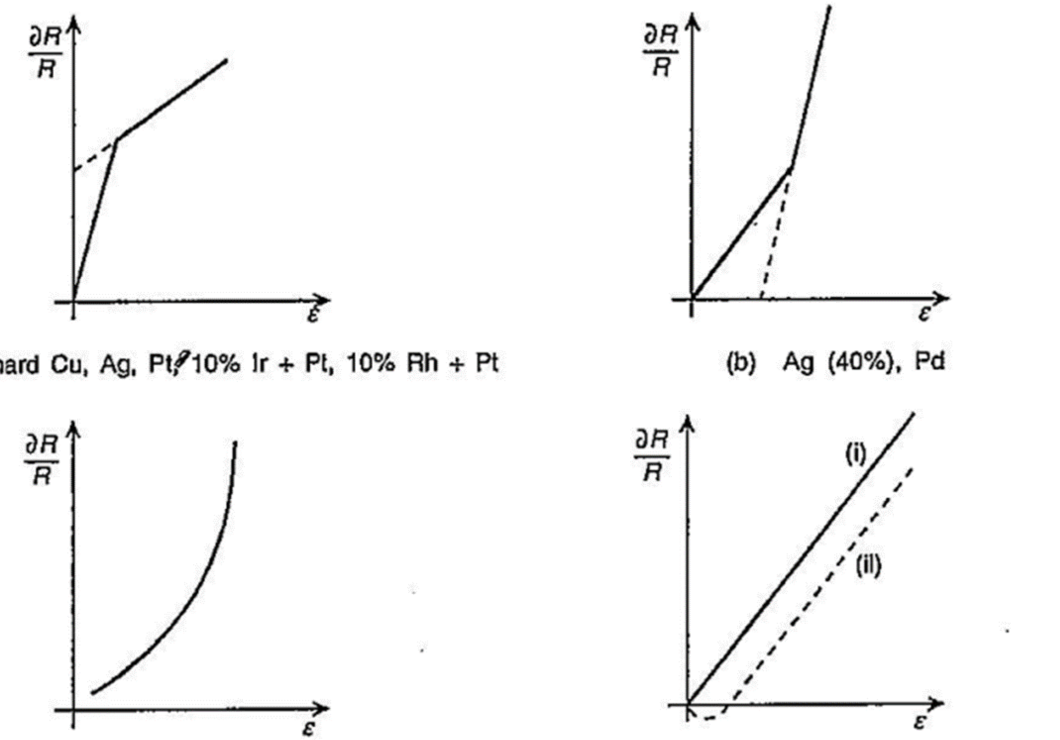
$$\lambda = \frac{\Delta R/R}{\Delta l/l} = 1 + 2\mu + \frac{\Delta \rho/\rho}{\Delta l/l} \tag{2.12}$$

- The last term on the right hand side of Eq. (2.12) is due to piezoresistance effect or Bridgeman effect and is often expressed as

$$\frac{\Delta\rho/\rho}{\Delta l/l} = \psi E$$

where ψ is the Bridgeman or longitudinal piezoresistance coefficient and E is the modulus of elasticity.

ne.



(a) Fe, hard Cu, Ag, Pt, 10% Ir + Pt, 10% Rh + Pt

(b) Ag (40%), Pd

(c) Minalpha

(d) (i) Annealed Cu/Ni, (ii) Hard drawn Ni

Fig. 2.8 Strain sensitivity for different materials.

- Unbonded strain gauges are used in preloaded conditions not to allow the 'strings' to go slack. The wires are nickel alloys such as Cu-Ni, Cr-Ni, or Ni-Fe with gauge factor between 2 and 4 and diameters varying from 0.02-0.03 mm.
- The bonded strain gauges are of a few types. When wire is used, the possibilities are

(i) Flat grid type,

(ii) Wrap around type, and

(iii) Woven type, although the flat grid type is more popular out of all the three.

- Etched foil type resistance strain gauge is one variety that, in recent years, has most extensively been used.

A gauge consists of the resistance element of proper design/shape, the gauge backing, cement, connection leads, and often protective coating or other protective means.

Flat Grid Bonded Strain Gauge

- The construction of the flat grid bonded strain gauge is shown in Fig. 2.9.
- Such construction has the advantage of better strain transmission from the member to the wire grid, small hysteresis and creep, and is more accurate when the strain member is thin.

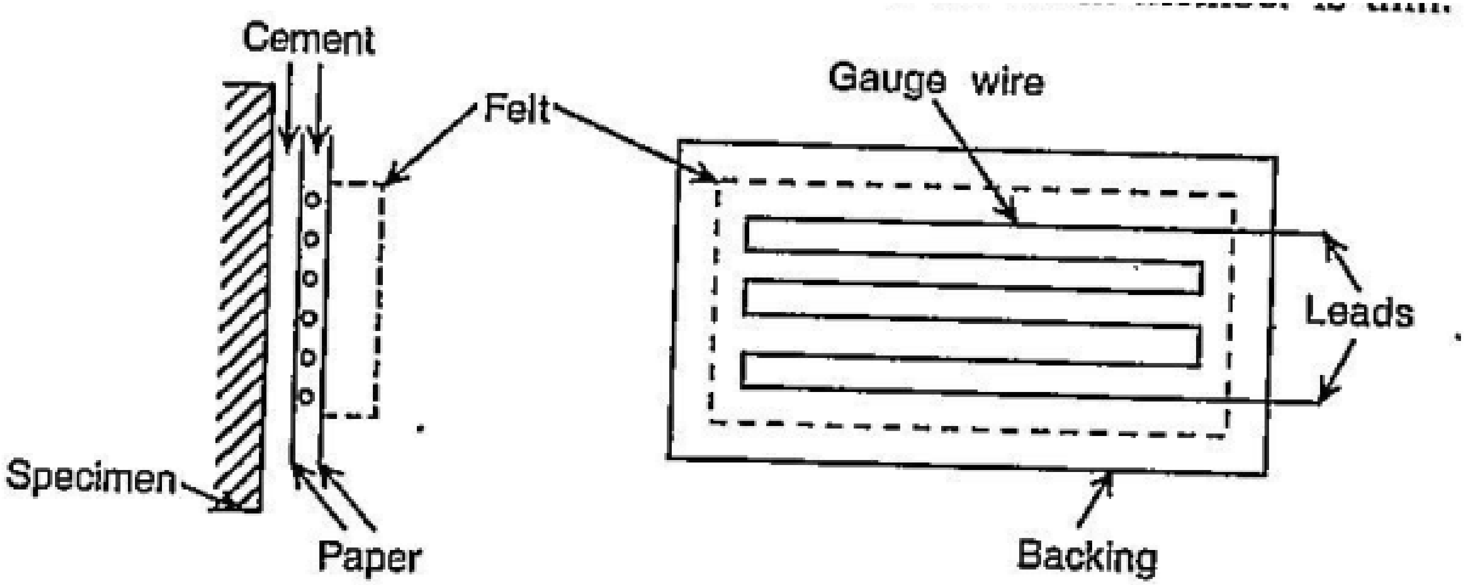


Fig. 2.9 Grid type gauge.

- The foil gauges are etched out from deposited films or sheets and have higher surface area to cross-section ratio than wire gauges, and hence, have better heat transfer property so that they can handle higher current.
- For wire gauges, the wires are usually drawn and often annealed, while bonded foil gauges consist of sensing elements which are formed from sheets of thickness less than 5×10^{-4} cm by photoetching processes so that any arbitrary shape can be given to these elements.
- Because the wire grid in the grid type structure has a finite width, the gauge has a sensitivity to transverse strain which may be as large as 2% of the longitudinal sensitivity. In foil grid structure, the end turns can be made wider or fat enough so that the transverse strain sensitivity is lesser. A typical grid structure foil gauge is shown in Fig. 2.10.

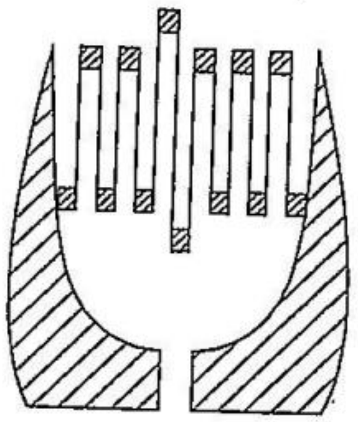


Fig. 2.10 Grid structure gauge with reduced transverse strain sensitivity.

- Vacuum deposition and sputter deposition thin film gauges are produced where cement between the elastic element and the gauge is not necessary for bonding.
- In the former case, a suitable elastic metal element which can be adapted for strain generation such as a diaphragm for pressure measurement, is placed in a vacuum chamber with a suitable dielectric material of much lower vapour point than the metal.
- With application of requisite amount of heat, this dielectric material vapourizes and then condenses and finally forms a thin layer on the metallic member.
- A template of a suitable shape is now placed over it and the evaporation-deposition process is repeated with the gauge material. Thus, the gauge is formed over the insulator substrate.

- In the sputtering-deposition process,
- The first step is nearly the same to form an insulating layer on the strain member.
- In the second step, without using a template, the metallic gauge material is sputtered over the entire substrate and the gauge pattern is defined by using micro-imaging techniques and photosensitive masking materials from outside the chamber and finally sputter-etching is used to remove all unmasked layers inside the vacuum chamber.

One important aspect of resistance gauges is its temperature coefficient of resistance.

The temperature at which the strain is measured may be different from the temperature when the strain member is bonded.

This gives rise to a differential expansion between the strain member and the gauge resulting in error in strain measurement.

Table 2.2 Strain gauge materials and their properties

Material	*Approx. nominal composition (%)*	*Gauge factor*	*Thermal coefficient of resistance (%/°C)*	*Nominal resistivity (μΩ cm)*
Constantan, Advance, Ferry	Ni 45, Cu 55	2.1–2.2	2×10^{-3}	0.45–0.48
Karma	Ni 74, Cr 20, Fe 3 Cu 3	2.1	2×10^{-3}	1.25
Nichrome V	Ni 80, Cr 20	2.2–2.6	10^{-2}	1.00
Isoelastic	Ni 36, Cr 8, Fe 52, Mn–Si–Mo 4	3.5–3.6	1.75×10^{-2}	1.05
Pt–W alloy	Pt 92, W 8	3.6–4.5	2.4×10^{-2}	0.62
Nickel	Ni 100	12	0.68	0.65
Manganin	Cu 84, Mn 12, Ni 4	0.3–0.48	2×10^{-3}	—
Platinum	Pt 100	4.8	0.4	0.1

The adhesives used to bond the gauge (backings) to the elastic member to be strained should be carefully selected. They must

(a) transmit the strain fully from the member surface to the gauge,
(b) have high insulation property,
(c) have high mechanical strength,
(d) have low thermal insulation,
(e) be as thin as possible yet provide strong bonding, and
(f) be suited to the environment, specifically the metal–paper and metal–dielectric interfacing.

Table 2.3 gives the properties of a few adhesives specially made for bonding strain gauges.

Table 2.3 Properties of adhesives

Material-base	*Temperature range* (°C)	*Cure-time* (hrs)	*Cure pressure* kg/cm^2	*Max. strain at room temp.* (%)	*Recommended lifetime* (yrs)
Acrylic	–75–65	1/12	Normal	10–15	1/2
Nitrocellulose	–75–65	24–48	1/2–1	10–15	2
Epoxy	0–200	12–24	1–3	6	1
Epoxy-phenolic	0–220	2	2–3	3–4	1
Polyimide	0–400	2–3	2.5–3	2–3	1/3
Ceramic	0–700	1	—	1/2	1

Acrylic has long term instability, nitrocellulose is a general purpose adhesive. Epoxy is resistant to moisture and has long term stability while epoxy-phenolic can be used in a thinner layer than the others. Polyimide and ceramic-base cements can be used at high temperatures though the latter is not very commonly used.

The recommended value of electrical insulation is of the order of 10^9 ohm at 50 V dc. If this value is not complied with, the gauge is likely to be 'shorted' and reading is susceptible to error. Most of the adhesives are vulnerable to high temperature and moisture/humidity which deteriorate their insulating as well as mechanical properties. Epoxy-base adhesives have been produced in various combinations with resins and hardeners for improving their properties.

Other than the adhesives given in Table 2.3, flame-spray and welding techniques have also been developed and are specifically used in some cases of free filament wire gauges. In the flame-spray, a solid rod is atomized to produce a ceramic spray which solidifies on the wires of the strain gauge making a bond without damaging the gauge or the strain member. This can be used upto about 800°C from near absolute zero while in the welding technique, the gauge is first epoxied to a thin metal shim. With low energy spot welder, the shim is then attached to the specimen. The foil gauges are specifically suitable with shim of thickness varying between 0.1–0.12 mm.

Gauges are made available in combinations often called 'rosettes' and these are designed in various configurations for specific stress–strain analysis and/or for transducer applications. A number of gauges are given relative orientations following certain pattern for the purpose. Thus, a three-gauge rosette used in stress analysis solves problems of a surface stress in magnitude and direction. Since the stress/strain is necessary to be measured at a point, it is best to stack these three gauges to form a rosette on that point. In fact, this sandwich pattern rosette is available from the manufacturers under the name 'stacked rosette'. Figure 2.11(a) shows such a three element rosette stacked at 45° to each other. In this, the topmost gauge is farthest from the specimen and all the gauges are insulated from each other, the topmost gauge gets heated up more compared to the bottommost which use the specimen as the heat sink. Two element stack type design is also commercially available. Such a design has an advantage that the strain/stress at the same point is sensed by all the gauges.

The alternative to the stack type design is the planar design which covers a small area rather than a point. Rosettes with such a design are available in two element 90° planar—usual and shear, three element 45°, 60° planar. They can be generated on the specimen as well. Figures 2.11(b), (c), and (d) show some of the types.

In fact, the technology of generating gauges on the specimen itself or on substrate as mentioned earlier by vacuum process has lead to wide scope of gauge pattern variation. It can be of any type depending on the specific requirements. The number of gauges at a location can also be changed as per this practice. Figure 2.11(e) shows a gauge pattern variation for measurement of strain in a diaphragm. Gauges 1 and 3 are subjected to tensile tangential stress while gauges 2 and 4 are subjected to compressive radial stress. 1′, 2′, 3′, and 4′ are contact terminals.

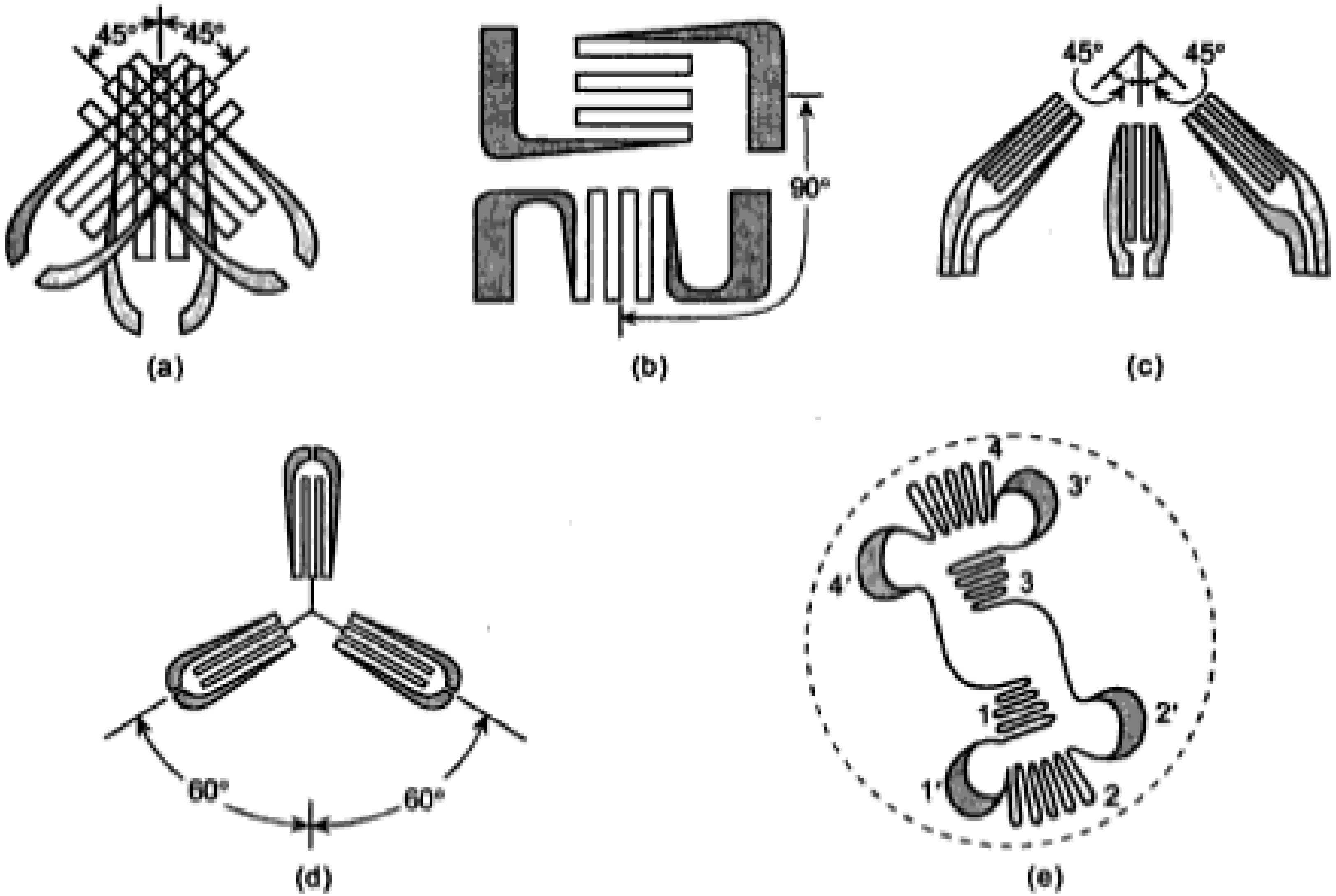


Fig. 2.11 Rosetted strain gauges: (a) three-element stacked type, (b) two-element right-angled, (c) three elements at 45° to each other, (d) three elements at 120° to each other, (e) gauge pattern on a diaphragm.

SEMICONDUCTOR STRAIN GAUGES

- First lot of semiconductor strain gauges were produced early in mid-thirties from single crystal silicon or germanium by cutting thin strips.
- It is still being done on the improvement of their performance and manufacturing ease because it has been known that although the semiconductor gauges have higher gauge factors, they are much inferior to the resistance types in so far as linearity and temperature stability are concerned.
- The discovery of semiconductor strain gauges has cleared the path of smart sensors, including production of strain sensitive cantilevers and diaphragms by doping selected small areas of monolithic silicon slice. Semiconductor strain gauges can be divided into two classes:

(i) Bonded semiconductor

(ii) Diffused semiconductor depending on their implementation.

- Strain sensitivity of semiconductor material depends, among other things, on the crystal material such as Si or Ge, doping levels (if any), type of doping materials, crystal cut-axis orientation, and so on.
- The bandgaps of both intrinsic and extrinsic semiconductors are affected by temperature variations.
- For intrinsic semiconductors, gauge factors are larger decreasing with increasing degrees of dopping, the thermal coefficients of resistivity also decreases.

As has been shown, the gauge factor of strain gauge is given by the relation

$$\lambda = 1 + 2\mu + \psi E$$

- The strain sensitivity of a semiconductor gauge is high and the large value is due to the large value of ΨE $(\Delta\rho/\rho)/(\Delta l/l)$, ,that is, specifically Ψ. The value of Poisson's ratio for semiconductors is less than that of metals although it is more in Si than Ge.
- There are a number of piezoresistive coefficients in a semiconductor material, they are called 'fundamental'.

- The longitudinal piezoresistive coefficients. in which the stress and current are in the same direction and the transverse piezoresistive coefficients, in which the stress and current are perpendicular to each other, are computed from these fundamental coefficients and the direction cosines of the current with respect to the crystallographic axes.

Table 2.4 Properties of semiconductor gauges

Material with crystal orientation	μ	E (10^{10} N/m^2)	ρ (10^{-3} Ωm)	λ (*longitudinal*)	*Thermal coefficient of resistance* β (10^{-5}/°C)
p-Si (111)	0.180	18.7	78	175	$70 \le \beta \le 700$
n-Si (100)	0.275	13.0	118	−135	$70 \le \beta \le 700$
p-Ge (111)	0.155	15.5	150	105	$70 \le \beta \le 700$
n-Ge (111)	0.156	15.5	160	−155	$70 \le \beta \le 700$

- Practical aspect of using a semiconductor strain gauge is governed by χ, R, gauge length, encapsulation/backing, bonding, leads geometry, and means of temperature compensation.
- Size and shape of the gauge are equally important. Some possible and useful shapes are given in Figs. 2.12(a), (b). (c). and (d). Sizes are determined by the specimen size as also resistance value R.

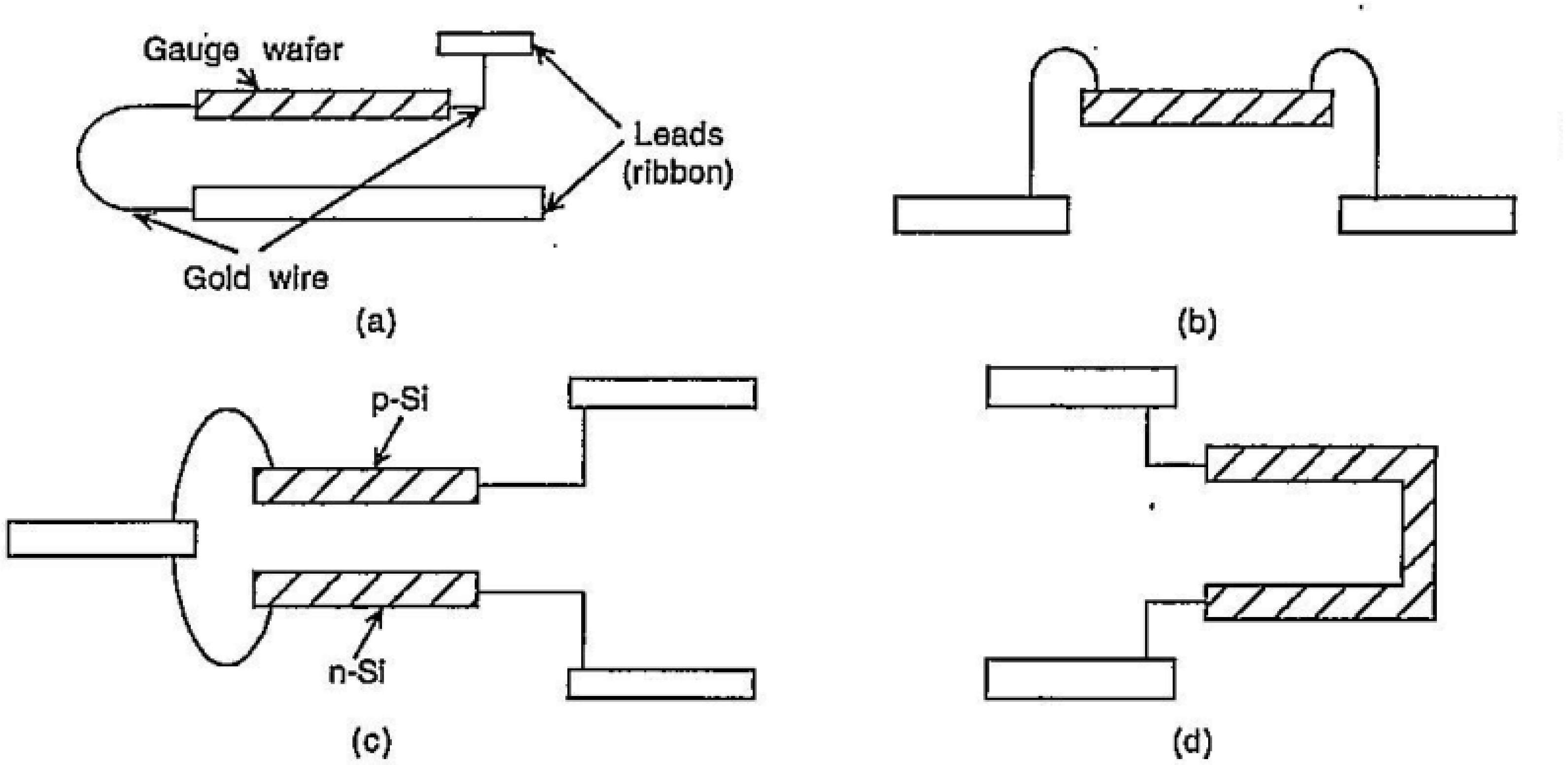


Fig. 2.12 Semiconductor gauges of different shapes and mountings.

- Semiconductor gauges with/without backing are bonded to the specimen with epoxy-based adhesives, or for better, diffused semiconductor gauges arc attached to the specimen by semiconductor diffusion process.
- The gauge is diffused directly on to the surface of the specimen such as a diaphragm, using photolithographic masking technique and an impurity such as boron is diffused into it.
- No separate bonding is necessary here. In recent times, the specimen, that is the strained member such as a cantilever or a diaphragm itself is also made from Si and the whole unit is developed into a smart sensor.
- A diaphragm of 2.5-25 mm diameter or cantilever of appropriate size is obtained in the main substrate of Si which is 50-750 mm in diameter.
- The four arm bridge is developed on this diaphragm as also the circuit of measurement by diffusion process.

The semiconductor strain gauge is basically nonlinear and an empirical relation between $\Delta R/R$ and ε

$$\frac{\Delta R}{R} = \sum_{j=1}^{n} k_j \varepsilon^j \tag{2.15}$$

is suggested, where k_j's are constants that depend on the materials and doping levels. Also, at high stress conditions temperature dependence of these coefficients are observed. Nonlinearity has been found to be improved by heavily doping the basic material of lower resistivity but then strain sensitivity is less. Often approximation by truncating the series upto $j = 2$ is good enough for practical use. Thus, an n-Si gauge of $\rho = 3.1 \times 10^{-4}$ ohm m would have

$$\lambda = -110 + 10^5 \varepsilon$$

and a p-Si with $\rho = 0.2 \times 10^{-3}$ ohm m would have

$$\lambda = 120 + 4 \times 10^4 \varepsilon$$

However, with higher resistivity such as $\rho = 78 \times 10^{-3}$ ohm m, a p-Si has a gauge factor (see Table 2.4)

$$\lambda = 175 + 7.26 \times 10^4 \varepsilon$$

As has been mentioned already, increasing doping decreases sensitivity towards temperature as well. Figure 2.13 shows the temperature–gauge factor curves for varying degrees of doping of a semiconductor gauge.

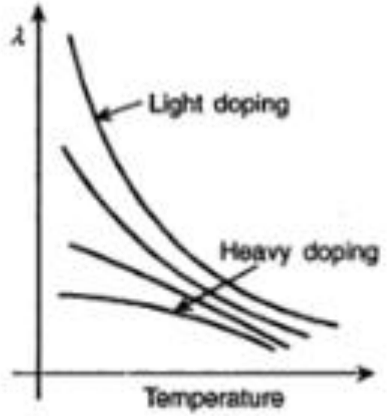


Fig. 2.13 Gauge factor versus temperature plots for different doping levels.

- In fact, doping changes the gauge resistance as well, decreasing it with high doping level. Figure 2.14 shows the *p*-*T* characteristics with doping as a parameter.

- Figure shows that higher doping gives high value of resistivity *p* and *β* the temperature coefficient of resistance—positive as well as high. But this occurs only up to a certain temperature above which the material behaves as in intrinsic conduction mode with negative temperature coefficient also of a very high value.

- However, with heavy doping. *p* is moderate and *β* quite small, and this condition persists over a wider temperature range.

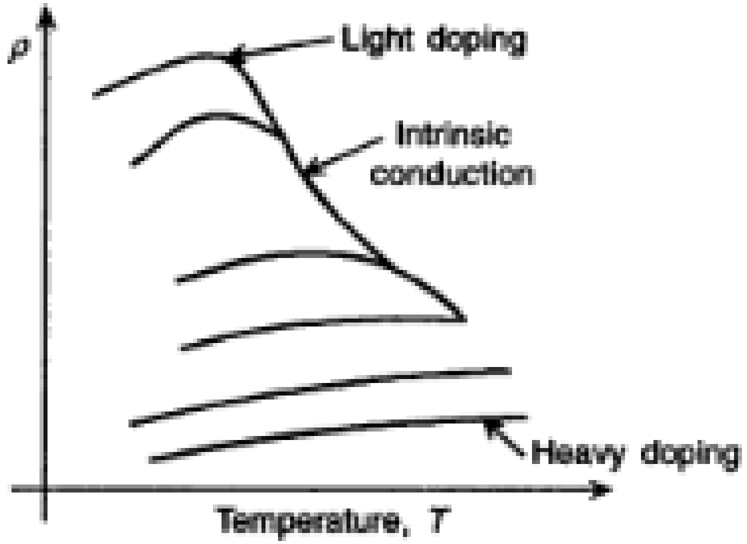


Fig. 2.14 Resistivity versus temperature for doping level variations.

- It has been discussed already, semiconductors under strain show piezoresistive effect which is so predominant over other effects related to Poisson's ratio and to on that based only on this dominant effect.
- Pressure transducers have been produced and within the elastic limits of silicon, electrical output is found proportional to mechanical strain or stress. The scheme consists of a cantilever beam of silicon about 0.1mm thick on to both sides of which planar resistors are produced by diffusion.
- Figure 2.15 shows the scheme with the header with connecting terminals. With the beam under stress, the resistors on the two sides of the beam undergo different changes because of compression on one side and extension on the other.
- The difference is measured by a bridge. The length l can be inserted in a pressure all where a diaphragm actuated by the inlet pressure is so mounted and attached to the cantilever that the deformation of the diaphragm is transmitted to the cantilever and hence, to the diffused resistance gauges.

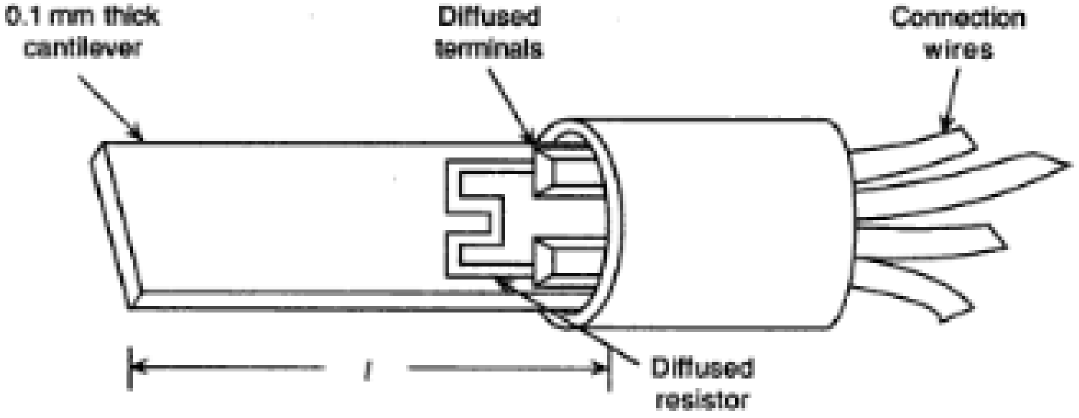


Fig. 2.15 Sensor using semiconductor piezoresistive effect.

For pressure measurement, thin silicon diaphragms with diffused resistors have been developed. A typical scheme is shown in Fig. 2.16. The piezoresistors are usually embedded in the diaphragm so that they get the strain of the diaphragm unabated.

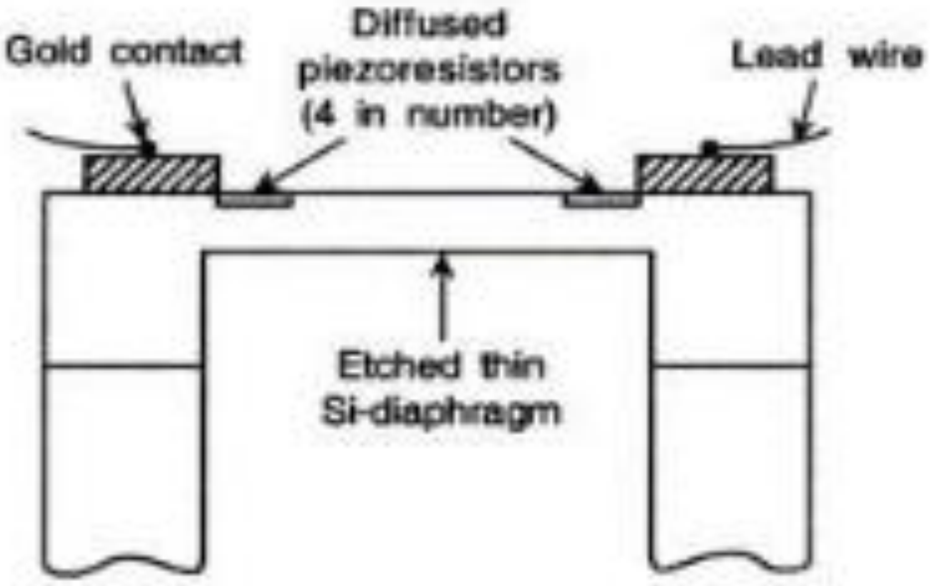


Fig. 2.16 Pressure measurement scheme using semiconductor diaphragm.

SESSION 9

- RESISTIVE TRANSDUCER: RTD, MATERIALS USED IN RTD.
- THERMISTOR: THERMISTOR MATERIAL, SHAPE.

RESISTIVE TRANSDUCER

- A resistive transducer is an electronic device that is capable of measuring various physical quantities like temperature, pressure, vibration, force, etc.
- The resistance of this transducer changes concerning the change in the physical quantities.
- These transducers can function in both primary as well as in secondary mode but most of the time it is used as secondary.
- The resistive transducers are of different types like resistive pressure transducers, thermistors resistors, LDR, etc.

Resistive Transducer (Contd.)

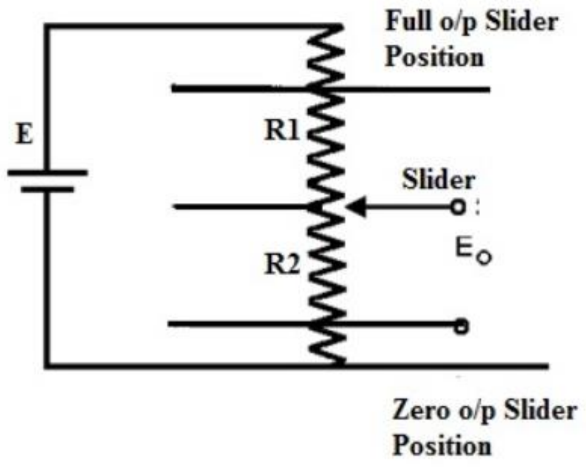
Working

- The working of a resistive transducer can be explained by considering a conductor rod as the transducer.
- The transducer works on the principle of the length of the conductor.
- The length of the conductor is directly proportional to its resistance and is inversely proportional to its cross-sectional area.
- Here, if we consider the length of the conductor as L, the cross-sectional area as A, the resistance as R and the resistivity as ρ, then the resistance can be denoted as
  - $R = {}^{\rho l}/_{A}$
- The resistance of the transducers can vary because of the change in environmental conditions as well as the physical properties of the conductor.
- Measuring devices like AC or DC can be used to measure the change in resistance.

Resistive Transducer (Contd.)

Resistive Transducer Circuit

- The resistive transducer consists of a long conductor whose length can be varied with time.
- One end of the conductor is connected while the other end is connected to a brush or a slider that can freely move along the length of the transducers.
- We can calculate the distance of the object by connecting the object to the slider of the resistive transducer.
- Whenever we apply energy to the object to displace it from its initial position, the slider will move along the length of the conductor as a result of which the length will change.



Resistive Transducer (Contd.)
Advantages AND DISADVANTAGES

- **Advantages**
- The resistive transducer can be used to give very quick results.
- The resistive transducers are available in various sizes and they have a considerably high amount of resistance.
- We can use both AC or DC for calculating the change in resistance.
- They are quite affordable and can be easily available in the market.
- W can use this transducer in various applications even when they are not a necessity.
- It can be used to give accurate results.
- **Disadvantages**
- A lot of power is wasted in moving the sliding contacts.
- The sliding contacts can produce a lot of noise.

Resistive Transducer (Contd.)

Applications

- A resistive transducer is mainly used to measure the temperature in various kinds of applications. When there is a change in temperature, the temperature coefficient of the resistive transducer changes which can be used to determine the change in temperature.
- The resistive transducer can function as a potentiometer where the resistance of the transducer can be varied by changing the length of the conductor.
- A resistive transducer can be used in the calculation of the displacement. When we apply strain on the resistor, the resistance changes. This characteristic can be used in the measurement of displacement, force, and pressure.

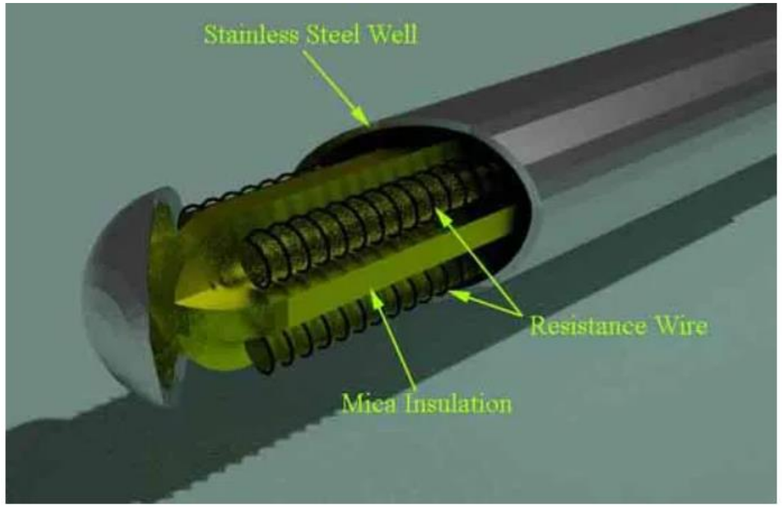
Resistance Temperature Detector (RTD)

- A Resistance Temperature Detector (also known as a Resistance Thermometer or RTD) is an electronic device used to determine the temperature by measuring the resistance of an electrical wire.
- This wire is referred to as a temperature sensor.
- The variation of resistance of the metal with the variation of the temperature is given as
    - $R_t = R_o\left[1 + (t - t_0) + \beta(t - t_0)^2 + \cdots\right]$
- Where, R_t and R_0 are the resistance values at t^oC and $t_0{}^oC$ temperatures. α and β are the constants depends on the metals.
- This expression is for huge range of temperature. For small range of temperature, the expression can be,
    - $R_t = R_o[1 + (t - t_0)]$

Resistance Temperature Detector (RTD)
Construction

- The construction is typically such that the wire is wound on a form (in a coil) on notched mica cross frame to achieve small size, improving the thermal conductivity to decrease the response time and a high rate of heat transfer is obtained.
- In the industrial RTD's, the coil is protected by a stainless steel sheath or a protective tube.
- The physical strain is negligible as the wire expands and increase the length of wire with the temperature change.
- If the strain on the wire is increasing, then the tension increases. Due to that, the resistance of the wire will change which is undesirable. So, we don't want to change the resistance of wire by any other unwanted changes except the temperature changes.
- Mica is placed in between the steel sheath and resistance wire for better electrical insulation.
- Due less strain in resistance wire, it should be carefully wound over mica sheet.

Resistance Temperature Detector (RTD)
Signal Conditioning



- In RTD the change in resistance value is very small with respect to the temperature.
- So, the RTD value is measured by using a bridge circuit.

By supplying the constant electric current to the bridge circuit and measuring the resulting voltage drop across the resistor, the RTD resistance can be calculated.

Temperature is determined by converting the RTD resistance value using a calibration expression.

Resistance Temperature Detector (RTD)

Two wire rtd bridge

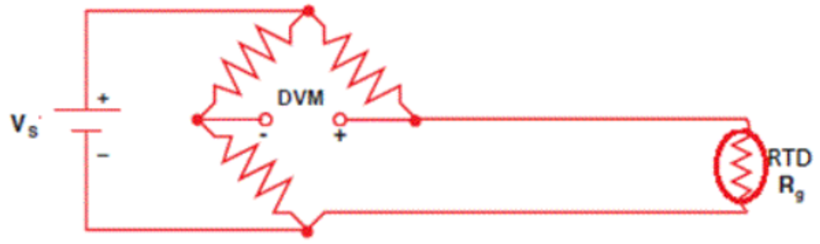
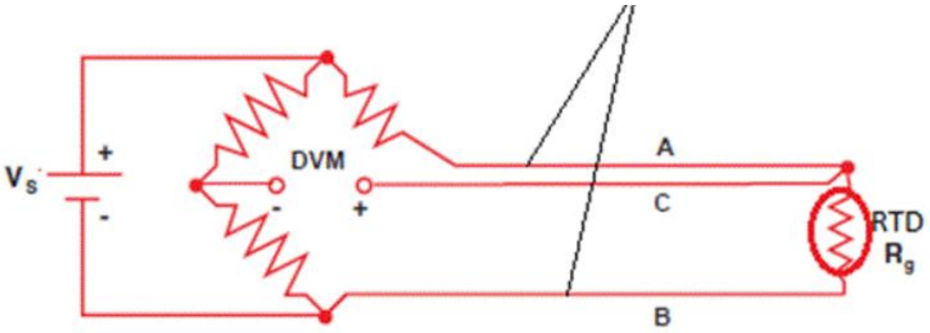


Fig.3. Two wires RTD Bridge

- In two wires RTD Bridge, the dummy wire is absent.
- The output taken from the remaining two ends as shown in fig
- If wires A and B are matched properly in terms of length and cross section area, then their impedance effects will cancel because each wire is in opposite position.
- In order to overcome that 3 wire RTD is introduced, which is done by adding a dummy wire

Resistance Temperature Detector (RTD)

THREE wire rtd bridge



- If wires A and B are matched properly in terms of length and cross section area, then their impedance effects will cancel because each wire is in opposite position.
- So that, the dummy wire C acts as a sense lead to measure the voltage drop across the RTD resistance and it carries no current
- In these circuits, the output voltage is directly proportional to the temperature. So, we need one calibration equation to find the temperature.

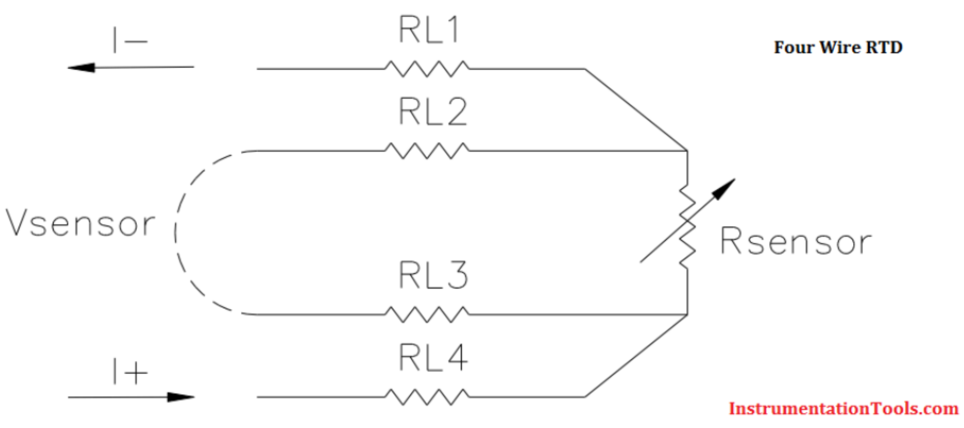
Resistance Temperature Detector (RTD)

4 Wire RTD Principle

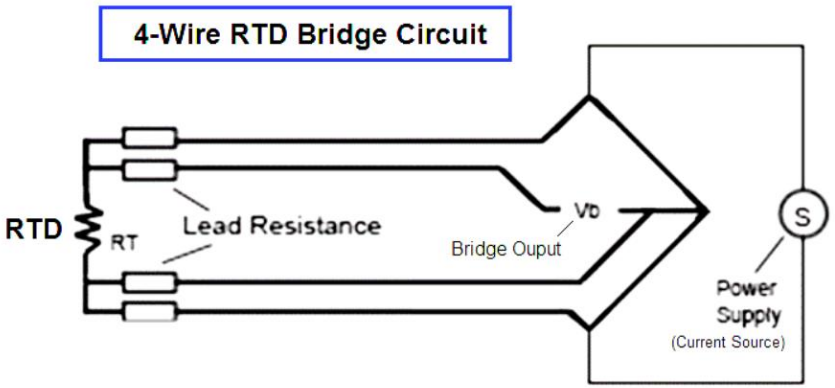
- The most accurate lead wire configuration is the “true” 4-wire configuration.
- In a true 4- wire configuration, the resistance of the lead wires does not contribute to the resistance of the sensor.
- 4-wire construction is used primarily where close accuracy is required.
- The true 4-wire measurement uses the current-potential method.
- A current of known value (I+) is passed through the sensor along the “current” lead wires.
- The voltage generated across the sensor is measured using the “potential” lead wires (Vsensor) and the sensor’s resistance is calculated by dividing the measured voltage by the Known current.

Resistance Temperature Detector (RTD)

4 Wire RTD Principle



- The 4-wire circuit is a true 4-wire bridge, which works by using wires 1 & 4 to power the circuit and wires 2 & 3 to read. This true bridge method will compensate for any differences in lead wire resistances.
- The resistance of the lead wires is not a factor because:
  - The value of the current is equal at any point in the circuit. It is independent of the resistance of the lead wire.
  - The input impedance of the voltage measurement circuitry is high enough to prevent any significant current flow in the voltage leads. Since no current is flowing, the voltage along the potential leads does not change along their length.



Resistance Temperature Detector (RTD)
materials used in RTD

- **Platinum Resistance Temperature Detectors**
  - Platinum RTDs are the most common type of RTD used in industrial applications.
  - This is because platinum has excellent corrosion resistance, excellent long-term stability, and measures a wide range of temperature, (-200…+850°C)
- **Nickel Resistance Temperature Detectors**
  - Nickel RTDs are less expensive than platinum and have good corrosion resistance.
  - However, nickel ages more rapidly over time and loses accuracy at higher temperatures. Nickel is limited to a measurement range of -80…+260°C.
- **Copper Resistance Temperature Detectors**
  - Copper RTDs have the best resistance to temperature linearity of the three RTD types, and copper is a low cost material.
  - However, copper oxidizes at higher temperatures. Copper is limited to a measurement range of -200…+260°C

Resistance Temperature Detector (RTD)
materials used in RTD

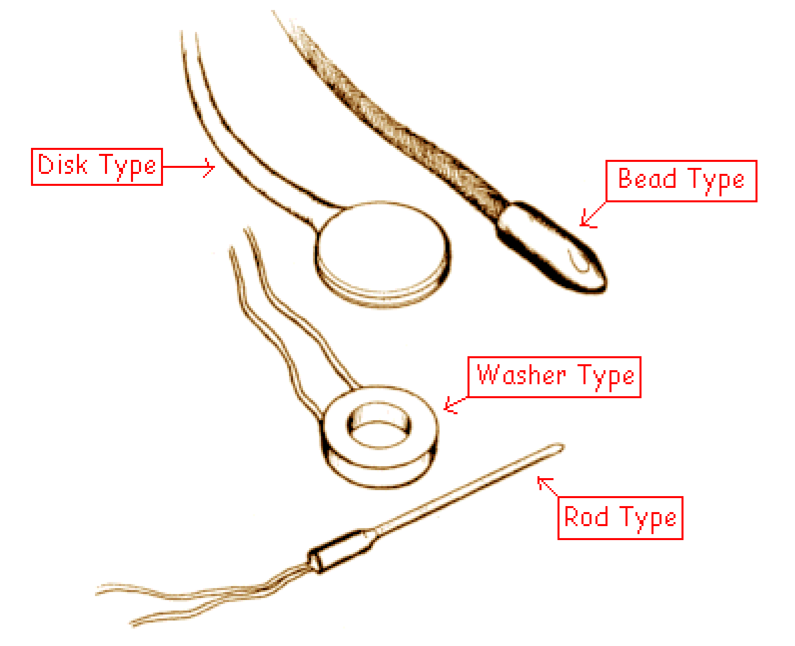
RTD type	Maximum measurement range	Long term stability	Corrosion resistance	Temperature vs. resistance linearity	Typical resistance at 0°C	Typical resistance at 100°C	Change in resistance 0...100°C	Resistance ratio $(R_{100}-R_0)/R_0$	Alpha (α) $(R_{100}-R_0)/(100 \times R_0)$
Platinum	-200...850°C	Excellent	Excellent	Good	100 Ω	138.5 Ω	38.5 Ω	0.385	0.00385
Nickel	-80...260°C	Fair	Good	Fair	120 Ω	200.64 Ω	80.64 Ω	0.672	0.00672
Copper	-200...260°C	Good	Fair	Excellent	9.035 Ω	12.897 Ω	3.86 Ω	0.427	0.00427

Resistance Temperature Detector (RTD)
ADVANTAGES AND DISADVANTAGES

- **Advantages:**
  - Good sensitivity
  - Linear over wide operating range
  - Uses standard copper wire
  - High temperature operating range
  - Copper RTD’s minimise thermocouple effect
  - Interchangeability over wide range
  - Works in wide temperature ranges
- **Disadvantages:**
  - Bulky in size and fragile
  - Slow thermal response time due to bulk
  - Self-heating problems
  - More susceptible to electrical noise
  - More expensive to test and diagnose

Thermistor

- The ‘Thermistor’ uses resistance to detect temperature.
- Thermistors can measure temperatures across the range of -40 ~ 150 ±0.35 °C (-40 ~ 302 ±0.63 °F).
- Typical operation resistances are in the kW range, although the actual resistance may range from few W to several MW.
- The adjoining figure shows typical types of thermistors.
- The shape of the thermistor probe can take the form of a bead, washer, disk, or rod.
- Basically, thermistors are broadly classified as Ceramic, PTC (positive temperature coefficient) and NTC (negative temperature coefficient) thermistors.



Typical Thermistor Types

** Image Courtesy - Engineering Fundamentals (www.efunda.com)*

**modified by Siddharth Rao*

Thermistor
Basic Working Principle

- The electrical resistance of metals depends on temperature.
- By measuring the changing resistance, the temperature can be determined.
- The change in resistance can easily be converted to an electrical signal transmittable.
- A thermistor is made of semiconductor, a mixture of metal oxide.
- Metals usually have a positive resistance coefficient with respect to temperature.
- Unlike metals, the semiconductors have a negative resistance coefficient.
- This is the main difference between a thermometer and a thermistor.
- Thus, it can be said that a PTC Thermistor is similar to an Resistance Temperature Detectors (RTD).

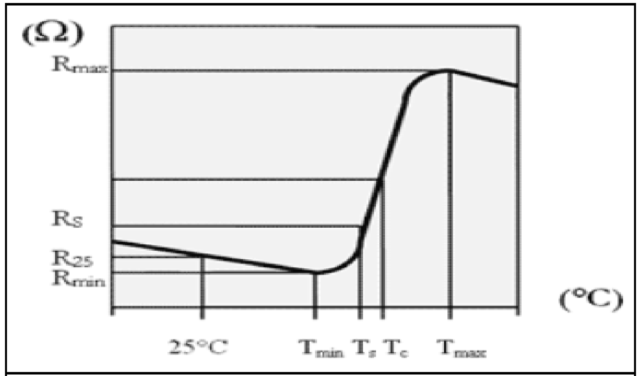
Thermistor
Basic Working Principle (Contd.)

- Thus, thermistors are based on the principle of when the temperature of the resistors changes, the electrical resistance of the resistors will change correspondingly.
- In Negative Temperature Coefficient (NTC) thermistors, when the temperature of the resistors increases, the resistance of the resistors will be decreased.
- In Positive Temperature Coefficient (PTC) thermistors, when the temperature of the resistors increases, the resistance of the resistors will also be increased.

Thermistor

Basic Working Principle - PTC

- The PTC (Positive Temperature Coefficient) is a temperature sensitive semiconductor, which is made of doped polycrystalline ceramic on the basis of barium titanate.
- The resistance of these thermistors increases sharply when a defined temperature is reached.
- This property is the reason for the self-regulation characteristic, which the PTC heating elements make use of.
- Due to the special Resistance-Temperature-characteristic, there is no additional temperature regulation or safety device necessary while reaching high heat-power level when using the low resistance area.
- The PTC-heating element regulates the power sensitively according to the required temperature. The power input depends on the requested heat output



The PTC - Working Principle

Image Courtesy - Dr. Baurecht GmbH - Kerms - Austria (www.ptc-ceramics.com)

**modified by Siddharth Rao*

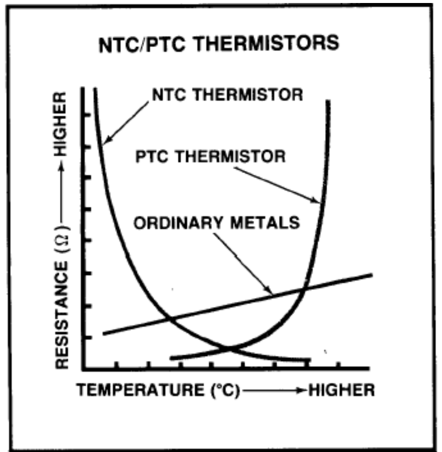
R	=	Resistance (Ohm)
R_{min}	=	Min. Resistance
R_{max}	=	Max. Resistance
R_s	=	Switch Resistance
R_{25}	=	Resistance at 25°C
T	=	Temperature (°C)
T_s	=	Switch Temperature
T_c	=	Curie Temperature
T_{min}	=	Temperature corresponding to R_{min}
T_{max}	=	Temperature corresponding to R_{max}

Image Courtesy - Dr. Baurecht GmbH - Kerms - Austria (www.ptc-ceramics.com)

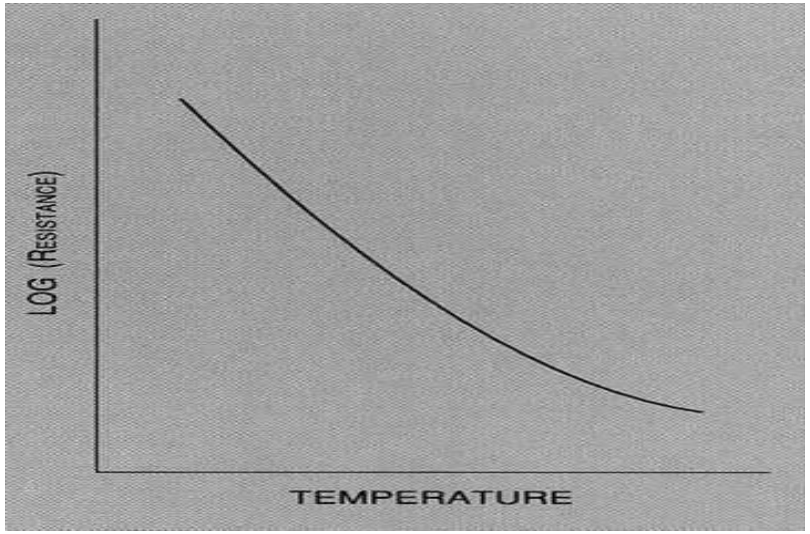
**modified by Siddharth Rao*

Thermistor
Basic Working Principle - NTC

- The NTC thermistors which are discussed herein are composed of metal oxides.
- The most commonly used oxides are those of manganese, nickel, cobalt, iron, copper and titanium.
- As seen from the adjoining figures, the resistance of these thermistors decreases with the increase in temperature.
- In the basic process of fabrication, a mixture of two or more metal oxide powders are combined with suitable binders, formed to a desired geometry, dried, and sintered at an elevated temperature.
- By varying the types of oxides used, their relative proportions, the sintering atmosphere, and the sintering temperature, a wide range of resistivities and temperature coefficient characteristics can be obtained.



The NTC / PTC Curves
Image Courtesy - www.autashap101.com/farms/h3.pdf
*modified by Siddharth Rao



Thermistor
Application

- The thermistor is a versatile component that can be used in a wide variety of applications where the measurand is temperature dependent.
- Depending on the type of application and the specific out put requirements, the PTC or the NTC Thermistor is used.
- Thus, the application have to be broadly divided as PTC Thermistor applications and NTC Thermistor applications respectively.
- Following are the various applications.

Thermistor
Application - PTC

- Over-current protection.
- Telecommunication applications.
- Picture tube degaussing.
- time delay and switching applications.
- motor starting.
- heating elements.
- ‘Fuse’ for Short-circuit and over-current protection.
- ‘switch’ for Motor start Degaussing.
- ‘temperature sensor’ in measurement and control & over temperature protection circuits.
- They are used to limit temperature for motor protection and over temperature protection circuits.
- They are also used as ‘level sensors’ and ‘limit indicators’.

Thermistor
Application - NTC

NTC thermistors are used in General Industrial Applications such as Industrial process controls, Photographic processing, Copy machines, Soldering irons (controlled), Solar energy equipment, etc

They are used in Consumer / Household Appliances like Thermostats, Burglar alarm detectors, Refrigeration and air conditioning, Fire detection, etc

They are used in Medical Applications like Fever thermometers, Dialysis equipment, Rectal temperature monitoring, Respiration rate measurement, Blood analysis equipment, Respirators, etc.

They are used in Instrumentation Applications like Motor winding compensation, Infrared sensing compensation, Instrument winding compensation, etc.

They are used in Automotive and Transportation Applications for Emission controls, Differential temperature controls, Engine temperatures, Aircraft temperatures, Rotor/bearing temperatures, etc

Thermistor
Application - NTC

- They are used in High Reliability Applications for monitoring Missiles & spacecraft temperatures, Aircraft temperature, Submarines & underwater monitoring and as a Fire control equipment.
- They are used in Communications Applications for Transistor temperature compensation, Gain stabilization, Piezoelectric temperature compensation.
- Apart from all these, they are also used in RF / Microwave power measurement, Voltage regulation circuits, Time delay devices, Sequential switching, Surge suppression, Inrush current limiting, etc.
- They are used in Food Handling Applications like Fast food processing, Perishable shipping, Oven temperature control, Coffee makers, Freezing point studies.

Thermistor
Advantages

- High accuracy, ~±0.02 °C (±0.36°F), better than RTDs, much better than thermocouples.
- High sensitivity, ~10 times better than RTDs, much better than thermocouples. As a result, lead wire and self-heating errors are negligible.
- Small in size compared to thermocouples.
- Response time shorter than RTDs, about the same as thermocouples.
- Reasonable long term stability and repeatability.

Thermistor
Limitations

- Limited temperature range, typically -100 ~ 150 °C (-148 ~ 302 °F).
- Nonlinear resistance-temperature relationship, unlike RTDs which have a very linear relationship.
- They can be affected by self-heating errors that result from excitation current being dissipated in the thermistor.
- Thermistors are also relatively fragile, so they must be handled and mounted carefully to avoid damage.
- Exposure to higher temperatures can de-calibrate a thermistor permanently, producing measurement inaccuracies.

Thermistor
SHAPES