Unit 4 – Sensors

Introduction – piezoelectric pressure sensor - capacitance pressure sensor - Capacitor plate sensor- piezoelectric devices for motion sensing - Hall effect based speed sensor – photodiodes –phototransistors - photovoltaic devices - Pyroelectric detector - semiconductor based IR sensors.

4.1. Introduction

A sensor is a device that measures a physical quantity and converts it into a signal which can be read by an observer or by an instrument. For example, a mercury-inglass thermometer converts the measured temperature into expansion and contraction of a liquid which can be read on a calibrated glass tube. A thermocouple converts temperature to an output voltage which can be read by a voltmeter. For accuracy, all sensors need to be calibrated against known standards.

Sensors are used in everyday objects such as touch-sensitive elevator buttons (tactile sensor) and lamps which dim or brighten by touching the base. There are also innumerable applications for sensors of which most people are never aware.

Applications include cars, machines, aerospace, medicine, manufacturing and robotics

A sensor is a device which receives and responds to a signal or stimulus. Here, the term "stimulus" means a property or a quantity that needs to be converted into electrical form. Hence, sensor can be defined as a device which receives a signal and converts it into electrical form which can be further used for electronic devices. A sensor differs from a transducer in the way that a transducer converts one form of energy into other form whereas a sensor converts the received signal into electrical form only

A sensor's sensitivity indicates how much the sensor's output changes when the measured quantity changes. For instance, if the mercury in a thermometer moves 1 cm when the temperature changes by 1 °C, the sensitivity is 1 cm/°C. Sensors that measure very small changes must have very high sensitivities. Sensors also have an impact on what they measure; for instance, a room temperature thermometer inserted into a hot cup of liquid cools the liquid while the liquid heats the thermometer. Sensors need to be designed to have a small effect on what is measured, making the sensor smaller often improves this and may introduce other advantages. Technological progress allows more and more sensors to be manufactured on a microscopic scale as microsensors using MEMS technology. In most cases, a microsensor reaches a significantly higher speed and sensitivity compared with Macroscopic approaches .

To illustrate the place of sensors in a larger system, Fig. 1.2 shows a block diagram of a data acquisition and control device. An object can be anything: a car, space ship, animal or human, liquid, or gas. Any material object may become a subject of some kind of a measurement. Data are collected from an object by a number of sensors.

Some of them (2, 3, and 4) are positioned directly on or inside the object. Sensor 1 perceives the object without a physical contact and, therefore, is called a *noncontact*

sensor. An example of such a sensor is a radiation detector and An example of such a sensor is a radiation detector and a TV camera. Even if we say ―noncontact‖, we remember that energy transfer always occurs between any sensor and an object.

Sensor 5 serves a different purpose. It monitors internal conditions of a data acquisition system itself. Some sensors (1 and 3) cannot be directly connected to standard electronic circuits because of inappropriate output signal formats. They require the use of interface devices (signal conditioners). Sensors 1, 2, 3, and 5 are passive. They generate electric signals without energy consumption from the electronic circuits. Sensor 4 is active. It requires an operating signal, which is provided by an excitation circuit. This signal is modified by the sensor in accordance with the converted information. An example of an active sensor is a thermistor, which is a temperature-sensitive resistor. It may operate with a constant-current source, which is an excitation circuit. Depending on the complexity of the system, the total number of sensors may vary from as little as one (a home thermostat) to many thousands (a space shuttle). Electrical signals from the sensors are fed into a multiplexer (MUX), which is a switch or a gate. Its function is to connect sensors one at a time to an analog-to-digital (A/D) converter if a sensor produces an analog signal, or directly to a computer if a sensor produces signals in a digital format. The computer controls a multiplexer and an A/D converter for the appropriate timing. Also, it may send control

signals to the actuator, which acts on the object. Examples of actuators are an electric motor, a solenoid, a relay, and a pneumatic valve. The system contains some peripheral devices (for instance, a data recorder, a display, an alarm, etc.) and a number of components, which are not shown in the block diagram. These may be filters, sample-and-hold circuits, amplifiers, and so forth.

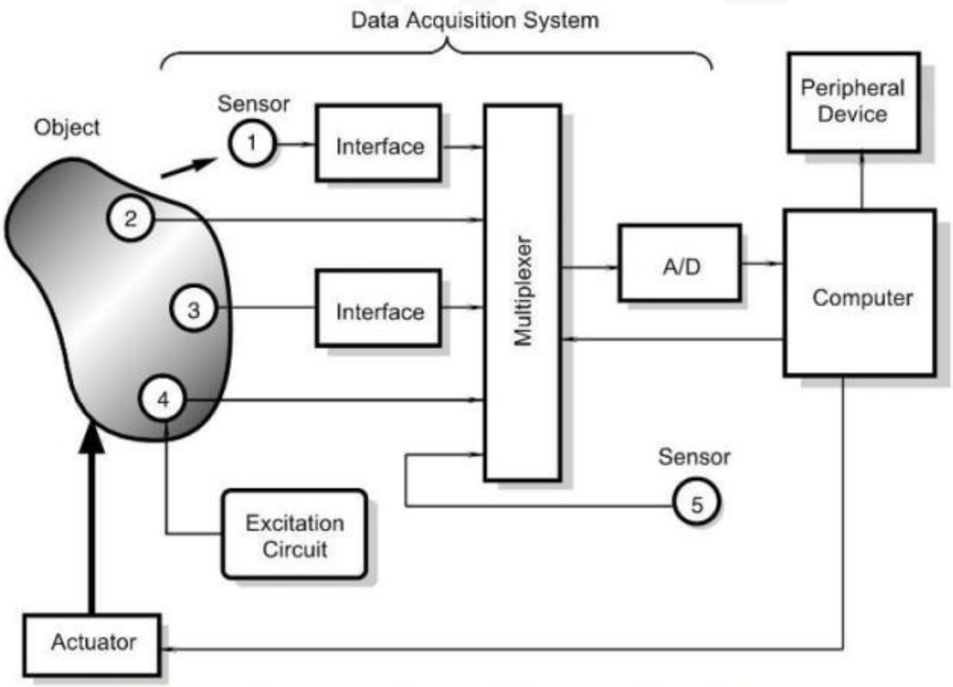


Figure (1.2): Positions of sensors in a data acquisition system. Sensor 1 is noncontact, sensors 2
and 3 are passive, sensor 4 is active, and sensor 5 is internal to a data acquisition system.

4.2 Classification of Sensors

All sensors may be of two kinds: **passive** and **active**. A passive sensor does not need any additional energy source and directly generates an electric signal in response to an external stimulus; that is, the input stimulus energy is converted by the sensor into the output signal.

The examples are a thermocouple, a photodiode, and a piezoelectric sensor. Most of passive sensors are direct sensors as we defined them earlier. The active sensors require external power for their operation, which is called an *excitation signal*. That signal is modified by the sensor to produce the output signal.

The active sensors sometimes are called *parametric* because their own properties change in response to an external effect and these properties can be subsequently converted into electric signals. It can be stated that a sensor's parameter modulates the excitation signal and that modulation carries information of the measured value.

For example, a thermistor is a temperature-sensitive resistor. It does not generate any electric signal, but by passing an electric current through it (excitation signal), its resistance can be measured by detecting variations in current and/or voltage across the thermistor. These variations (presented in ohms) directly relate to temperature through a known function. Another example of an active sensor is a resistive strain gauge in which electrical resistance relates to a strain. To measure the resistance of a sensor, electric current must be applied to it from an external power source.

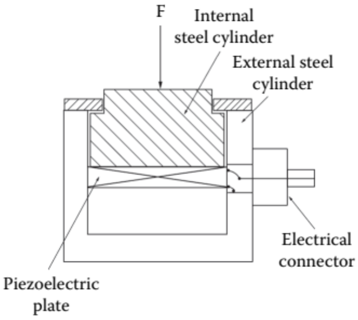
Depending on the selected reference, sensors can be classified into **absolute** and **relative**. An *absolute* sensor detects a stimulus in reference to an absolute physical scale that is independent on the measurement conditions, whereas a *relative* sensor produces a signal that relates to some special case. An example of an absolute sensor is a thermistor: a temperature-sensitive resistor. Its electrical resistance directly relates to the absolute temperature scale of Kelvin. Another very popular temperature sensor—a thermocouple—is a relative sensor. It produces an electric voltage that is function of a temperature gradient across the thermocouple wires. Thus, a thermocouple output signal cannot be related to any particular temperature without referencing to a known baseline. Another example of the absolute and relative sensors is a pressure sensor. An absolute-pressure sensor produces signal in reference to vacuum—an absolute zero on a pressure scale. A relative-pressure sensor produces signal with respect to a selected baseline that is not zero pressure (e.g., to the atmospheric pressure).Another way to look at a sensor is to consider all of its properties, such as what it measures (stimulus), what its specifications are, what physical phenomenon it is sensitive to, what conversion mechanism is employed, what material it is fabricated from, and what its field of application is.

Listing of common measured variables:

• Temperature, • Pressure, • Flow rate, • Composition and • Liquid level.

4.3. *Piezoelectric Pressure Sensor*

A basic piezoelectric pressure sensor cell is illustrated in Figure (a). The piezoelectric plate is mechanically connected to a pressure sensing diaphragm that acts as a force summing device and creates a stress in the plate. The diaphragm, in turn, is exposed to a gas pressure that acts through a perforated cap. The cap protects the sensor against dust, mechanical damage, and other hostile environmental conditions. A change in pressure causes a mechanical change in the piezoelectric plate, which results in a generation of electric charge or voltage, which is then transferred out of the cell using a shielded connector. A commercial model of a pressure sensor is shown in Figure (b). Once a sensor has been designed and constructed, various sensor parameters are usually supplied by the sensor manufacturer.



(a)

(b)

Many applications exist for the stress/pressure sensor. Examples of stress or force sensor applications include the measurement of compression, tension, impact, vibrating, balancing, striking, rolling, cutting, forming, pressing, machining, and punching operations. Examples for pressure sensor applications include gas pressure, combustion, explosion, pulsation, actuation, cavitation, fluidic, pneumatic, blast, turbulence, and sound pressure measurements.

Advantages:

1. Simple sensing mechanism, direct conversion of mechanical quantities into electrical quantities

2. Passive or self-generating sensors

3. High sensitivity

4. High-frequency response, i.e., for measuring of small pressure fluctuations; high mechanical stiffness, i.e., they undergo small deformations (~micron or angstrom) and hence do not load the measurand

5. Light weight

6. Small size (~1 mm)

7. Rugged

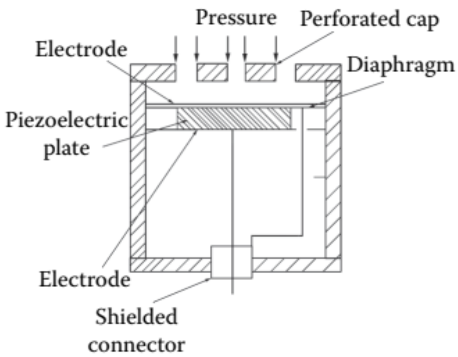
8. Nominally low cost

Disadvantages:

1. High-output impedance

2. Sensitive to temperature

3. Brittle



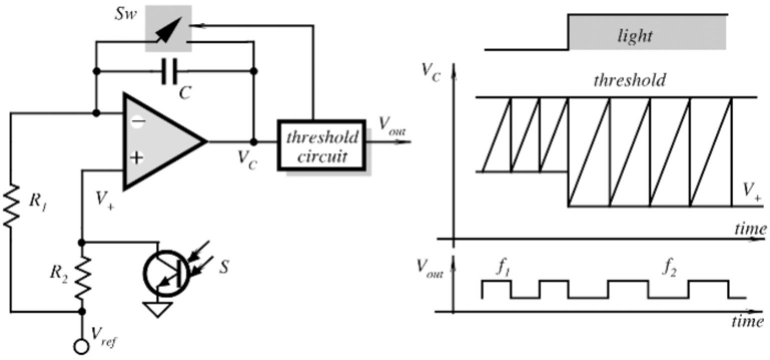
Cross section view of a basic sensor

4.4 Capacitive pressure sensor

A silicon diaphragm can be used with another pressure-to-electric output conversion process: in a capacitive sensor. Here, the diaphragm displacement modulates capacitance with respect to the reference plate (backplate). This conversion is especially effective for the low-pressure sensors. An entire sensor can be fabricated from a solid piece of silicon, thus maximizing its operational stability. The diaphragm can be designed to produce up to 25% capacitance change over the full range which makes these sensors candidates for direct digitization. Whereas a piezoresistive diaphragm should be designed to maximize stress at its edges, the capacitive diaphragm utilizes a displacement of its central portion. These diaphragms can be protected against overpressure by including mechanical stops close to either side of the diaphragm (for a differential pressure sensor). Unfortunately, in the piezoresistive diaphragms, the same protection is not quite effective because of small operational displacements. As a result, the piezoresistive sensors typically have burst pressures of about 10 times the full-scale rating, whereas capacitive sensors with overpressure stops can handle 1000 times the rated full-scale pressure. This is especially important for the low-pressure applications, where relatively high-pressure pulses can occur.

Figure (A) shows a simplified circuit diagram of a modulating oscillator. It is composed of an integrator built with an operational amplifier and a threshold circuit. The voltage across the capacitor, C, is an integral of the current whose value is proportional to the voltage in the noninverting input of the operational amplifier.

When that voltage reaches the threshold, switch SW closes, thus fully discharging the capacitor. The capacitor starts integrating the current again until the cycle repeats. The operating point of the amplifier is defined by the resistor R_2, a phototransistor S, and the reference voltage Vref . A change in light flux which is incident on the base of the transistor changes its collector current, thus shifting the operation point. A similar circuit may be used for direct conversion of a resistive transducer, (e.g., a thermistor).



(A) (B)

While designing a capacitive pressure sensor, for good linearity it is important to maintain flatness of the diaphragm. Traditionally, these sensors are linear only over the displacements which are much less than their thickness. One way to improve the linear range is to make a diaphragm with groves and corrugations by applying micromachining technology. Planar diaphragms are generally considered more sensitive than the corrugated diaphragms with the same size and thickness. However, in the presence of the in-plane tensile stresses, the corrugations serve to release some of the stresses, thus resulting in better sensitivity and linearity. Capacitive sensors are very popular in many applications. Currently, micromachining technology allows us to fabricate small monolithic capacitive sensors.

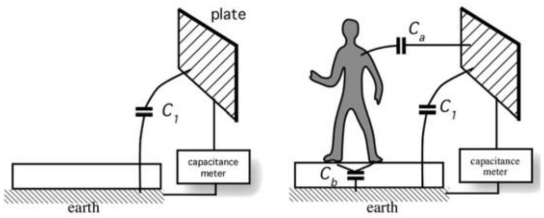
4.5. Capacitive plate sensor

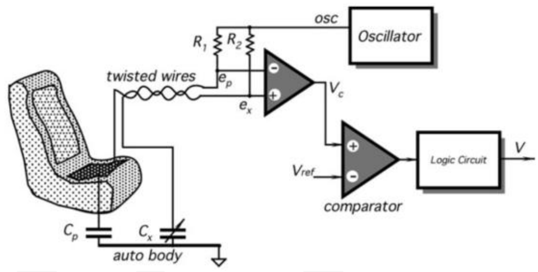
Figure 1 shows that the capacitance between a test plate and earth is equal to C_1. When a person moves into vicinity of the plate, it forms two additional capacitors: one between the plate and its own body, C_a, and the other between the body and the earth, C_b. Then, the resulting capacitance C between the plate and the earth becomes larger by the incremental capacitance DC.

$$C = C_1 + \Delta C = C_1 + \frac{C_a C_b}{C_a + C_b}. \quad (1)$$

With the appropriate apparatus, this phenomenon can be used for occupancy detection. What required is to measure capacitance between a test plate (the probe) and a reference plate (the earth). Figure 2 illustrates a capacitive security system for an automobile. A sensing probe is imbedded into a car seat. It can be fabricated as a metal plate, metal net, a conductive fabric, etc. The probe forms one plate of a capacitor C_p. The other plate of the capacitor is formed either by a body of an automobile, or by a separate plate positioned under a floor mat. A reference capacitor C_x is composed of a simple fixed or trimming capacitor, which should be placed close to the seat probe. The probe plate and the reference capacitor are, respectively, connected to two inputs of a charge detector (resistors R_1 and R_2). The conductors preferably should be twisted to reduce the introduction of spurious signals as much as possible.

For instance, strips of a twinflex cabling were found quite adequate. A differential charge detector is controlled by an oscillator, which produces square pulses (Fig. 3). Under a no-seat-occupied condition, the reference capacitor is adjusted to be approximately equal to C_p. Resistors and the corresponding capacitors define time constants of the networks. Both RC circuits have nearly equal time constants τ_1. Voltages across the resistors are fed into the inputs of a differential amplifier, whose output voltage V_c is near zero. Small spikes at the output is the result of some unavoidable imbalance. When a person is positioned on the seat, her body forms an additional capacitance in parallel with C_p, thus increasing a time constant of the R_1C_p-network from τ_1 to τ_2. This is indicated by the increased spike amplitudes at the output of a differential amplifier. The comparator compares V_c with a predetermined threshold voltage Vref. When the spikes exceed the threshold, the comparator sends an indication signal to the logic circuit that generates signal V manifesting the car occupancy. It should be noted that a capacitive detector is an active sensor, because it essentially required an oscillating test signal to measure the capacitance value.





4.6. Piezoelectric devices for motion sensing

Displacement Motion (*z*-Direction)

Parallel plate capacitive arrangement can be used as a motion sensor utilizing the spacing variation in the dielectric material when the spacing change is less than the electrode size. The motion detector is designed in such a way that one plate of the capacitor is fixed, while the other is movable perpendicular to the plane of the electrodes. The parallel plate capacitance shows that capacitance is inversely related to spacing *d*. Since the sensitivity of the sensor dC/dd is large for small *d* following the parabolic capacitance–motion displacement relationship, the sensor conveniently yields large variation in capacitance at small displacement due to motion. However, the nonlinearity of the sensor poses problems during calibration. The linearity in sensor during motion detection through displacement measurement can be solved by plotting the impedance *ZC*–displacement *d* curve which is linear following the relation $ZC = d/(j\omega\varepsilon 0\varepsilon rA)$.

Shear Motion (*x* Direction)

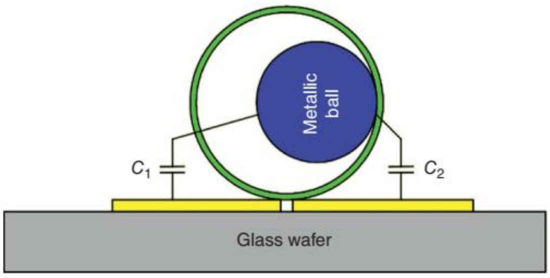
The shear motion detector utilizes the overlapping electrode of the parallel plate capacitor architectural design. The overlapping area principle is suitable when the displacement incurred due to motion is larger than the dimensions of the electrodes. Here, the transverse motion of the electrodes relative to the other changes the capacitance linearly with shear displacement. Quite long excursions are possible with good linearity, but the gap needs to be small and well controlled. As with spacing variation, overlap is needed so that unwanted sensitivities are minimized.

Tilt Sensor

Tilt sensors are devices that produce an electrical signal that varies with an angular movement. These sensors are used to measure slope and tilt within a limited range of motion. Sometimes, the tilt sensors are referred to as inclinometers. These devices consist of multiple pair of combs electrodes-based capacitors and central proof mass or other dielectric liquid. When a tilt occurs, the central mass/liquid moves toward one of the combs, so the capacitance increases at one side and decreases at the other side.

Rotary Motion Sensor

The rotary motion detector was designed by Lee et al. and consists of movable TE in the form of a metallic ball and separate and fixed BEs as shown in Figure. The top metallic ball and the BEs are separated by a plastic pipe along whose inner walls the metallic ball rotates through an angle of 360 . The sensor uses the spatial distance between the metallic ball and each of the BEs to determine the respective capacitance developed between them. The capacitances developed between the metallic ball and each of the BEs due to the spatial position of the metallic ball under tilt gives the tilt in the sensor.



4.7. Hall effect sensor

This effect was discovered by E. H. Hall in the year 1879. After his name, the effect is called the Hall effect. As per the effect, whenever a current carrying semiconductor or conductor block is kept under a perpendicularly aligned magnetic field regarding the electric current as illustrated in Figure 8.3, then the magnetic field applies a transverse force on the electric charges flowing through the medium and thus the charges start accumulating at one side of the conductor or semiconductor and thus a voltage is formed between two sides of the material block. The Hall effect is the name given to this phenomenon, and Hall voltage is the resultant voltage.

The fundamentals of the Hall effect are as follows. If a lengthy strip with current flowing through it is exposed to a dc magnetic field, the Hall effect manifests in its most basic and traditional form. The Lorentz force affected the carriers present in the strip.

$$F = e\,E + e\,v \times B \qquad (1)$$

Here, e stands for electrical charge of a carrier, E for local electrical field, v for charge carrier's velocity, and B for the magnetic flux density, which we assume to be parallel to the strip plane. Suppose the material of the strip is an N-type highly extrinsic semiconductor. We ignore the fact that there are holes. An external electrical field called Ee is applied in the x direction along the length of the strip. That external field is mostly responsible for the electrical field E. The electrical field outside affects how the electrons behave. The electrons move down the strip with the average drift velocity in response to the external electrical field, $v_{dn} = \mu_n.\ Ee$, μ_n being the drift mobility of electrons. The corresponding current density is provided by, $J_n = q.\ n.\ \mu_n.\ Ee$, (2)

Where q denotes the fundamental charge. The carrier velocity v is due to the thermal agitation and drift. Let us neglect for a moment the thermal motion. The magnetic component of the Lorentz force is then provided by, $F_{mn} = q\ \mu_n\ Ee \times B$ (3)

That force pushes the electrons toward the strip's upper edge. Due to this, the concentration of electron at the strip's top and lower edges, respectively, rises and falls. An electrical field emerges between the strip edges as a result of those space charges. This electrical field acts on the electrons by a force,

$$F_{en} = -\ q\ E_H \qquad (4)$$

That force tends to decrease the excess charges at the strip's edges. As a steady state, the two transverse forces F_{mn} and F_{en} balance. By equating both the equations of F_{mn} and F_{en}, we find

$E_H = \mu_n Ee \times B$ (5)

The transverse electrical field E_H is called the Hall electric field. Without neglecting the thermal agitation of the electrons, we get, $E_H = - \mu_n H_n [Ee \times B]$. (6)

Here, $\mu_n H_n$ denotes the Hall mobility of the electrons. The Hall mobility differs a little from the drift mobility: It is given by, $\mu H_n = r_H \mu_n$, where r_H is the Hall scattering factor. This numerical parameter captures the impact of the carriers' thermal motion and their scattering on the Hall effect. Another useful expression for the Hall electrical field is obtained when the external electrical field expressed before is expressed by the current density, $E_H = - R_H [J \times B]$. (7)

Here, RH denotes the Hall coefficient, in this case given by, $R_H = 1/q\ n$, , where n denotes the density of free electrons. Here, again we use the appropriate sign because we neglected the thermal agitation of charge carriers. Without neglecting their thermal agitation instead of we obtain, $R_H = r_H/(q\ n)$.

The development of a measurably high voltage between the strip's edges is the hallmark of the Hall effect. It is known as the Hall voltage. Let us choose two points on the opposite corners of the strip in Figure 8.3 so that their potential difference is zero when $B = 0$. The Hall voltage is then provided by,

$$V_{Hall} = \int_{S_2}^{S_1} E_H ds$$

In this particular case, we find that $V_{Hall} = \mu_n H_n Ee\ B\ w$, where w denotes the width of the strip. Once more, the strip's current density may be measured by, $J = I / t\ w$.

Here, I denote the current in the strip and t is the thickness of the strip. So, the source of Hall voltage is, $V_{Hall} = (R_H/t)\ I\ B$.

The above equations describe the basics of Hall effect. However, considering correction factors and other practical factors, the magnitude of Hall voltage increases with the increase in applied magnetic field B. Additionally, the current and voltage biasing, respectively, can be provided via,

$$V_{Hall} = G \frac{w}{l} \mu_H V_{Bias} = S_{V_v} V_{Bias} B \quad (8)$$

$$V_{Hall} = G \frac{w}{n \times e \times t} I_{Bias} B = S_{V_I} I_{Bias} B \quad (9)$$

Here, G = geometrical correction factor, μ_H = Hall mobility of majority carriers, r_H = the Hall scattering factor, V_{Bias} = total bias voltage, n = carrier concentration, e = electron charge, I_{Bias} = total bias current, t = thickness of the n-well implantation, B = external magnetic field, $S_{V\ V}$ = voltage-related voltage-mode sensitivity, and $S_{V\ I}$ = current-related voltage-mode sensitivity.

The voltage-related voltage-mode sensitivity ($S_{V\ V}$), which is influenced using the geometry and Hall mobility, is significantly reliant on temperature. Hence, $S_{V\ V}$ will change with temperature as well. On the other hand, the carrier concentration (n) has an inverse relationship with the current-related voltage-mode sensitivity (S_{VI}). This term is constant for most of the operating temperature and also for a plate doping density in the range of 10^{15} and 10^{17} cm^{-3}.

Selective material for a Hall sensor can increase the device performance and hence extremely significant in selecting a perfect material for the fabrication. According V_{Hall} equation, materials having increased mobility and decreased conductivity will work well as Hall effect sensor materials. Thus, metals which have increased conductivity and low mobility are not a good option for Hall sensors. The sensors are generally fabricated using n-type semiconductors since the majority carrier, i.e. electron, has higher mobility than holes. The suitable materials, in this case, are Si and III–V semiconductors like InSb, InAs, and GaAs. High mobility and significant conductivity are features of III–V semiconductors. Si is preferred to fabricate the Hall device despite having moderate mobility since it is compatible with IC technology. Table 8.2 describes different parameters such as the energy band gap (Eg), the carrier mobility (μ_n), and Hall coefficient (R_H) at 300 K of various semiconductors suitable for the fabrication of Hall plates. The Hall coefficient (R_H) in this case is also calculated for a fixed doping density.

The voltage mode operation of the Hall sensor is highly popular. In this case, the output is measured in terms of voltage. The current is the sensor's output in the case of a current mode Hall sensor though. However, the current mode is very seldom used in practice. The offset voltage or output voltage when zero magnetic field is applied is one of the key issues with Hall devices. The current mode Hall sensor has a similar construction to already available gadgets. The same option of offsetting the offset brought on by mismatch exists with the present mode device. The way in which signals are taken out makes the most impact.

The voltage mode Hall sensor is illustrated in Figure 8.4a, where the bias current I_{Bias} flows from D to B arm of the plate. In this case, a symmetric structure is fabricated for current spinning, which reduces the offset of the device. Hall voltage develops across the orthogonal arm (AC) under an applied magnetic field (BZ) in the plane's direction.

Figure 8.4b shows the current mode, where the current has been injected laterally in two adjacent arms (A and B) and an applied magnetic field causes unbalanced output current from the other two adjacent arms (C and D). The difference between the output current, in this case, can be characterized as a corresponding current source of a Hall current I_{Hall}. The principal of the same can be found in previous reports, where the Hall current is defined as,

$$I_{Hall} = \mu_H \frac{w}{l} B I_{bias} \quad (10)$$

where I_{Bias} = total bias current, μH = Hall mobility of majority carriers, w = width-to-length ratio of the plate, and Bl = normal magnetic field. To comprehend its use and application, this technique has undergone thorough study. The expression of output currents (I_{H+}, I_{H-}) can be given as,

$$I_{H^+} = \frac{I_{Bias}}{2} + \frac{I_{Hall}}{2} \quad (11)$$

$$I_{H^-} = \frac{I_{Bias}}{2} - \frac{I_{Hall}}{2} \quad (12)$$

In this case, the Hall current (I_{Hall}) is relational to the biasing current (I_{Bias}), external magnetic field (B_z), and magnetic resistance coefficient (β). The Hall plate current with a cross shape may be expressed as,

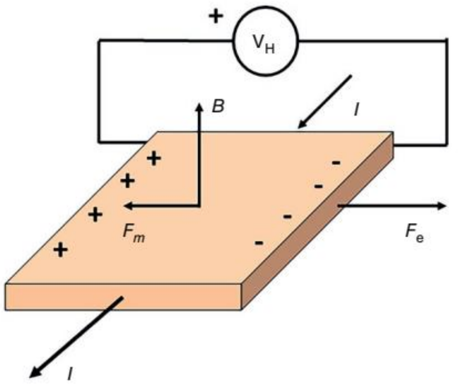
$$I_{Hall} = \mu_H \frac{\beta B_Z I_{bias}}{1-(\beta B_Z)^2} \quad (13)$$

Where β is calculated as $\beta = \frac{R_{(B_Z)} - R_{(B_{Z=0})}}{R_{(B_Z)} \times R_{(B_{Z=0})}}$ (14)

where $R(Bz = 0)$ and $R(Bz)$ represent the Hall plate resistance in the absence and presence of an external magnetic field, respectively. For a high mobility plate, the current mode Hall sensor offers superior resolution in frail magnetic fields. Additionally, the system may be made smaller by using fewer terminals thanks to the existing style of operation. The present mode is anticipated to receive greater attention in the near future due to the current technical trend of miniaturization.

This technique can be used to figure out the doping type of a semiconductor as the polarity of Hall voltage depends on the majority carrier present in the semiconductor. Hall effect has major applications in industries to manufacture low-power sensors, position detection, and contactless switching. These magnetic sensors are often high performance and cost-effective. These are also integrable with CMOS technologies. Although temperature has effects on the performance of different Hall sensors, temperature drift and offset management can provide optimal behavior and high sensitivity. Often sensor geometries are optimized in Hall sensor fabrications to get good results. However, reports have shown that the geometry change may create a significant offset variation. Several sensors for position detection, biomolecule detection, chemical detection, and gas detection have been reported using the Hall effect sensing mechanism.

Figure 8.3 The Hall effect experimental setup.



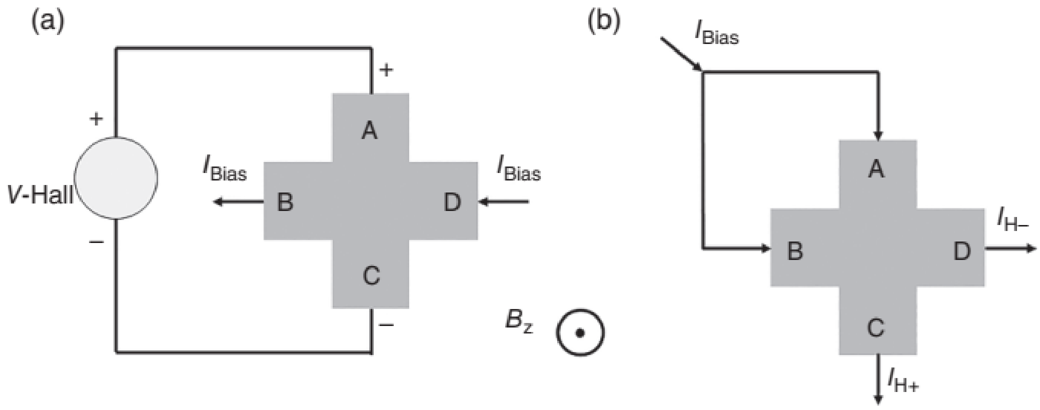


Figure 8.4 (a) Hall plate in voltage mode operation. (b) Hall plate in current mode operation.

4.8. Photodiodes

Photodiode is a solid-state device which converts incident light into an electric current. It is made of Silicon. It consists of a shallow diffused p-n junction, normally a p-on-n configuration. When photons of energy greater than 1.1eV (the bandgap of silicon) fall on the device, they are absorbed and electron-hole pairs are created. The depth at which the photons are absorbed depends upon their energy. The lower the energy of the photons, the deeper they are absorbed. Then the electron-hole pairs drift apart. When the minority carriers reach the junction, they are swept across by the electric field and an electric current establishes. Photodiodes are one of the types of photodetector, which convert light into either current or voltage. These are regular semiconductor diodes except that they may be either exposed to detect vacuum UV or X-rays or packaged with a opening or optical fiber connection to allow light to reach the sensitive part of the device.

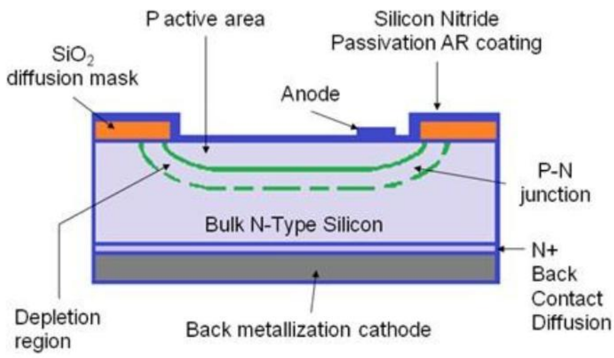


Figure shows the construction of Photo diode detector. It is constructed from single crystal silicon wafers. It is a p-n junction device. The upper layer is p layer. It is very thin and formed by thermal diffusion or ion implantation of doping material such as boron. Depletion region is narrow and is sandwiched between p layer and bulk n type layer of silicon. Light irradiates at front surface, anode, while the back surface is cathode. The incidence of light on anode generates a flow of electron across the p-n junction which is the measure of light intensity.

Applications of photo diodes

Camera: Light Meters, Automatic Shutter Control, Auto-focus, Photographic Flash Control **Medical:** CAT Scanners - X ray Detection, Pulse Oximeters, Blood Particle Analyzers **Industry:** Bar Code Scanners, Light Pens, Brightness Controls, Encoders, Position Sensors, Surveying Instruments, Copiers - Density of Toner.

4.9. Phototransistors

A photodiode directly converts photons into charge carriers, specifically one electron and one hole (hole–electron pair) per a photon. The phototransistors can do the same, and in addition to provide current gain, resulting in a much higher sensitivity. The collector–base junction is a reverse-biased diode. If the transistor is connected into a circuit containing a battery, a photoinduced current flows through the loop, which includes the base–emitter region. This current is amplified by the transistor in the same manner as in a conventional transistor, resulting in a significant increase in the collector current. The energy bands for the phototransistor are shown in Fig. 14.9. The photon induced base current is returned to the collector through the emitter and the external circuitry. In so doing, electrons are supplied to the base region by the emitter where they are pulled into the collector by the electric field. The sensitivity of a phototransistor is a function of the collector–base diode quantum efficiency and also of the dc current gain of the transistor. Therefore, the overall sensitivity is a function of collector current. When subjected to varying ambient temperature, collector current changes linearly with a positive slope of about 0.00667 per C. The magnitude of this temperature coefficient is primarily a result of the increase in current gain versus temperature, since the collector–base photocurrent temperature coefficient is only about 0.001 per C. The family of collector current versus collector voltage characteristics is very much similar to that of a conventional transistor. This implies that circuits with phototransistors can be designed by using the regular methods of transistor circuit techniques, except that its base should be used as an input of a photoinduced current that is supplied by its collector. Since the actual photogeneration of carriers occurs in the collector–base region, the larger the area of this region, the more carriers are generated; thus, the phototransistor is designed so to offer a large area to impinging light. A phototransistor can be either a two-lead or a three lead device. In the latter case, the base lead is available, and the transistor may be used as a standard bipolar transistor with or without the additional capability of sensing light, thus giving a designer greater flexibility in circuit development. However, a two-lead device is the most popular as a dedicated photosensor. When the base of the transistor is floating, it can be represented by an equivalent circuit shown in Fig. 14.10. Two capacitors Cc and Ce represent base–collector and base–emitter capacitances, which are the speed-limiting factors. Maximum frequency response of the phototransistor may be estimated from $f_l \sim g_m / 2C_e$ where f_l is the current-gain-bandwidth product and gm is the transistor's forward transconductance.

Whenever a higher sensitivity of a photodetector is required, especially if high response speed is not of a concern, an integrated Darlington detector is recommended. It is comprised of a phototransistor whose emitter is coupled to the base of a bipolar transistor. Since a Darlington connection gives current gain equal to a product of current gains of two transistors, the circuit proves to be an efficient way to make a sensitive detector.

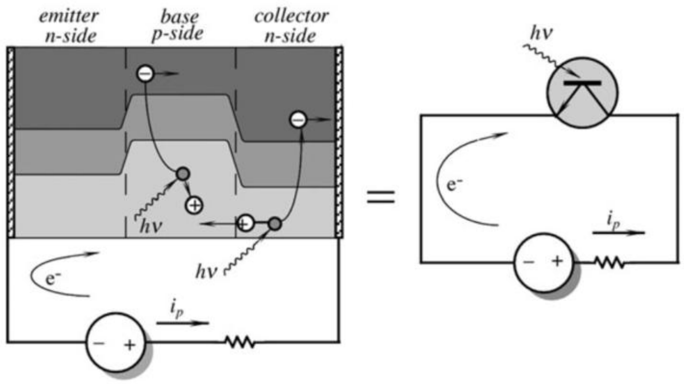
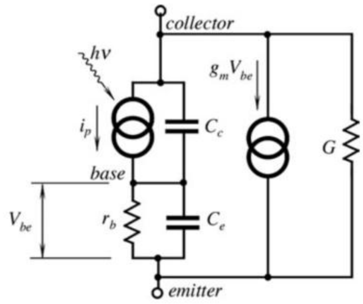


Fig. 14.9 Energy bands in a phototransistor

14 Light Detectors



4.10. Photovoltaic devices

Solid-state quantum detectors (photovoltaic and photoconductive devices) relay on the interaction of individual photons with a crystalline lattice of semiconductor materials. Their operations are based on the photo effect that was discovered by Albert Einstein, which won him the Nobel Prize. In 1905, he made a remarkable assumption about the nature of light that at least under certain circumstances, its energy was concentrated into localized bundles, later named photons. The energy of a single photon is given by $E = h\nu$ (1)

where ν is the frequency of light, and h = 6.626075x10^{-34} Js (or 4.135671015 eVs) is Planck's constant derived on the basis of the wave theory of light. When a photon strikes a surface of a conductor, it may result in the generation of a free electron. Part (Φ) of the photon energy E is used to detach the electron from the surface, while the other part gives to the electron its kinetic energy. The photoelectric effect can be described as $h\nu = \phi + K_m$ (2)

where Φ is called the work function of the emitting surface, and K_m is the maximum kinetic energy of the electron upon its exiting the surface. The similar processes occur when a semiconductor pn-junction is subjected to radiant energy: The photon transfers its energy to an electron and, if the energy is sufficiently high, the electron may become mobile, which results in an electric current. If energy is not sufficient for liberating an electron, the photon energy is just converted to heat.

The periodic lattice of crystalline materials establishes the allowed energy bands for electrons that exist within that solid. The energy of any electron within the pure material must be confined to one of these energy bands that may be separated by gaps or ranges of forbidden energies. That is, the electron can have only "permitted" energies.

Figure 14.1a shows energy bands of a semiconductor material, where E_g is the magnitude in eV of the forbidden band gap. The lower band is called the valence band, which corresponds to those electrons that are bound to specific lattice sites within the crystal. In the case of silicon or germanium, they are parts of the covalent bonding that constitute the interatomic forces within the crystal. The next higher lying band is called the conduction band and represents electrons that are free to migrate through the crystal. Electrons in this band contribute to the electrical conductivity of the material. The two bands are separated by the band gap, the size of which determines the whether the material is classified as a semiconductor or an isolator. The number of electrons within the crystal is just adequate to completely fill all available sites within the valence band. In the absence of thermal excitation, both isolators and semiconductors would therefore have a configuration in which the valence band is completely full, and the conduction band is completely empty.

Under these imaginable circumstances, neither would theoretically show any electrical conductivity. In a metal, the highest occupied energy band is not completely full. Therefore, electrons can easily migrate throughout the material. Metals are characterized by very high electrical conductivity. In isolators or semiconductors, on the other hand, the electron must first cross the energy band gap in order to reach the conduction band, and the conductivity is therefore many orders of magnitude lower. For isolators, the band gap is usually 5 eV or more, whereas for semiconductors, the gap is considerably less. Note that the longest the wavelength (lower frequency of a photon), the less energy is required to originate a photo effect.

When the photon of frequency n1 strikes the crystal, its energy is high enough to separate the electron from its site in the valence band and push it through the band gap into a conduction band at a higher energy level. In that band, the electron is free to serve as a current carrier. The deficiency of an electron in the valence band creates a hole that also serves as a current carrier. This is manifested in the reduction of specific resistivity of the material. On the other hand, Fig. 14.1b shows that a photon of lower frequency n2 does not have sufficient energy to push the electron through the band gap. The energy is released as heat without creating current carriers.

The energy gap serves as a threshold below which the material is not light sensitive. However, the threshold is not abrupt. Throughout the photon-excitation process, the law of conservation of momentum applies. The momentum and density of hole-electron sites are higher at the center of both the valence and conduction bands, and fall to zero at the upper and lower ends of the bands. Therefore, the probability of an excited valence-band electron finding a site of like momentum in the conduction band is greater at the center of the bands, and the lowest at the ends of the bands. Therefore, the response of a material to photon energy increases from E_g.

gradually to its maximum and then falls back to zero at the energy corresponding to the difference between the bottom of the valence band and the top of the conduction band. A typical spectral response of a semiconductive material is shown in Fig. 14.2. The light response of a bulk material can be altered by adding various impurities. They can be used to reshape and shift a spectral response of the material. All devices that directly convert photons of electromagnetic radiation into charge carriers are called quantum detectors, which are generally produced in the form of photodiodes, phototransistors, and photoresistors.

When comparing the characteristics of different photodetectors, the following specifications usually should be considered: Noise equivalent power (NEP) is the amount of light equivalent to the intrinsic noise level of the detector. Stated differently, it is the light level required to obtain a signal-to-noise ratio equal to unity. Since the noise level is proportional to the square root of the bandwidth, the NEP is expressed in units of W/Hz^{-2}

NEP = noise current / radiant sensitivity (λ_p)

D* refers to the detectivity of a detector's sensitive area of 1 cm^2 and a noise bandwidth of 1 Hz

$$D = \frac{\sqrt{area\,(cm)^2}}{NEP}$$

chopping frequency and the spectral content must be also specified. The detectivity is expressed in the units of Hz^{-2}/W. It can be said that the higher the value of D*, the better the detector.

IR cutoff wavelength (λ_c) represents the long-wavelength limit of spectral response, and often is listed as the wavelength at which the detectivity drops by 10% of the peak value.

Maximum current is specified for photoconductive detectors (such as HgCdTe) that operate at constant currents. The operating current never should exceed the maximum limit.

Maximum reverse voltage is specified for Ge and Si photodiodes and photoconductive cells. Exceeding this voltage can cause the breakdown and severe deterioration of the sensor's performance.

Radiant responsivity is the ratio of the output photocurrent (or output voltage) divided by the incident radiant power at a given wavelength, expressed in A/W or V/W.

Dark current ID for photodiodes is a leakage current at a reverse voltage when the diode is in complete darkness. This current generally is temperature dependent and may vary from pA to mA. It approximately doubles for every 10C increase in temperature.

Field of view (FOV) is the angular measure of the volume of space where the sensor can respond to the source of radiation.

Junction capacitance (C_j) is similar to the capacitance of a parallel plate capacitor. It should be considered whenever a high-speed response is required. The value of Cj drops with reverse bias and is higher for the larger diode areas

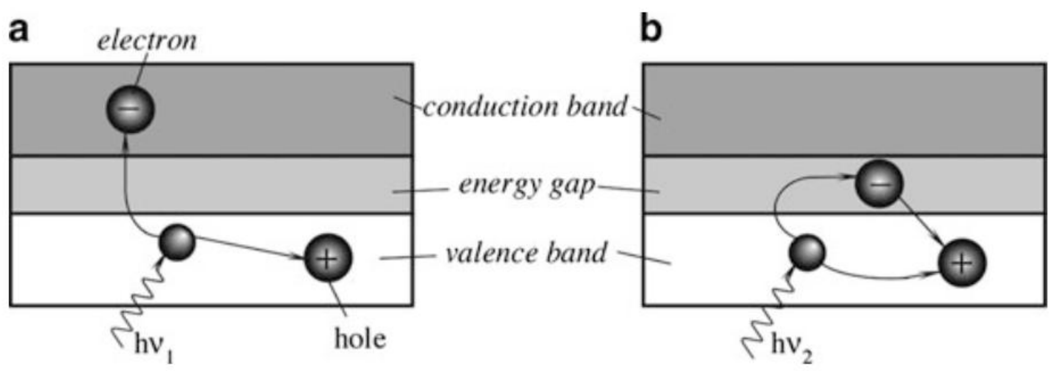


Fig. 14.1 Photoeffect in a semiconductor for high (**a**) and low (**b**) energy photons

4.11. Pyroelectric detector

Figure 3.27 depicts a pyroelectric detector (pyroelectric sensor) connected to a resistor R_b that represents either the internal leakage resistance or a combined input resistance of the interface circuit, which is connected to the sensor. The equivalent electrical circuit of the sensor is shown at right. It consists of three components: (1) the current source generating a heat induced current, i (remember that a current is a movement of electric charges), (2) the sensor's capacitance, C, and (3) the leakage resistance, R_b. The output signal from the pyroelectric sensor can be taken in form of either charge (current) or voltage, depending on the application. Being a capacitor, the pyroelectric device is discharged when connected to a resistor, R_b. Electric current through the resistor and voltage across the resistor represent the heat flow induced charge. It can be characterized by two pyroelectric coefficients

$P_q = \frac{dP_s}{dT}$ Pyroelectric charge coefficient (1)

$P_V = \frac{dE}{dT}$ Pyroelectric voltage coefficient (2)

where P s is the spontaneous polarization (which is the other way to say "electric charge"), E is the electric field strength, and T is the temperature in K. Both coefficients are related by way of the electric permittivity, ε_r and dielectric constant, ε_0.

$$\frac{P_Q}{P_V} = \frac{dP_s}{dE} = \varepsilon_r \varepsilon_0 \qquad (3)$$

The polarization is temperature dependent and, as a result, both pyroelectric coefficients (3) are also functions of temperature.

If a pyroelectric material is exposed to a heat source, its temperature rises by DT and the corresponding charge and voltage changes can be described by the following equations:

$$\Delta Q = P_Q A \Delta T \qquad (4)$$

$$\Delta V = P_V h \Delta T \qquad (5)$$

Remembering that the sensor's capacitance can be defined as

$$C_e = \frac{\Delta Q}{\Delta V} = \varepsilon_r \varepsilon_0 \frac{A}{h} \qquad (6)$$

From (4) and (6)

$$\Delta V = P_Q \frac{A}{C_e} \Delta T = P_Q \frac{\varepsilon_r \varepsilon_0}{h} \Delta T \qquad (7)$$

It is seen that the peak output voltage is proportional to the sensor's temperature rise and pyroelectric charge coefficient and inversely proportional to its thickness. When the pyroelectric sensor is subjected to a thermal gradient its polarization (electric charge developed across the crystal) varies with the temperature of the crystal. A typical polarization-temperature curve is shown in Fig. 3.28. The voltage pyroelectric coefficient, Pv, is a slope of the polarization curve. It increases dramatically near the Curie temperature at which the polarization disappears, and the material permanently loses its pyroelectric properties. The curves imply that the sensor's sensitivity increases with temperature at the expense of nonlinearity.

Piezo- and pyroelectric materials such as lithium tantalate and polarized ceramics are typical materials to produce the pyroelectric sensors. During recent years, a deposition of pyroelectric thin films have been intensively researched. Especially promising is use of lead titanate ($PbTiO_3$), which is a ferroelectric ceramic having both a high pyroelectric coefficient and a high Curie temperature of about 490C. This material can be easily deposited on silicon substrates by the so called sol-gel spin casting deposition method

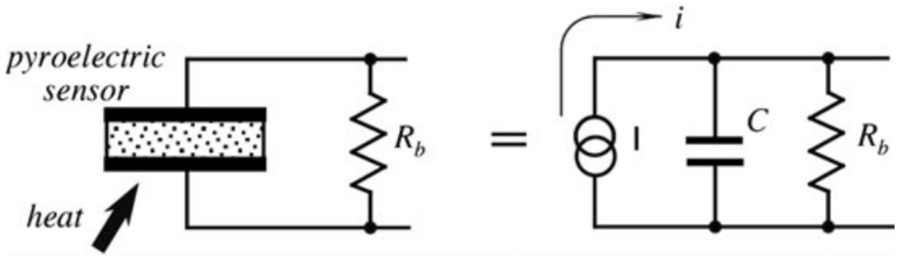


Fig. 3.27 Pyroelectric sensor and its equivalent circuit

Figure 3.29 shows the timing diagrams for a pyroelectric sensor when it is exposed to a step function of heat. It is seen that the electric charge reaches its peak value almost instantaneously, and then decays with a thermal time constant, τ_T. The physical meaning is this: a thermally induced polarization occurs initially in the most outer layer of the crystalline material (just few atomic layers), whose temperature nearly instantaneously raises to its maximum level. This creates the highest thermal gradient across the material thickness, leading to the maximum polarization. Then, heat propagates through the material, being absorbed by its mass in proportion to a thermal capacitance, CT, and some of the heat is lost to the surroundings through a thermal resistance, RT. This diminishes the initial gradient that generates the electric charge. The thermal time constant is a product of the sensors' thermal capacity and thermal resistance $\tau_T = C_T R_T = cAhR_T$ where c is the specific heat of the pyroelectric element. The thermal resistance RT is function of all thermal losses to the surroundings through convection, conduction, and thermal radiation. For the low-frequency applications, it is desirable to use sensors with τ_T as large as practical, while for the high-speed applications (for instance, to measure laser pulses), a thermal time constant should be dramatically reduced. For that purpose, the pyroelectric material may be laminated with a heat sink: a piece of aluminum or copper. When a pyroelectric sensor is exposed to a heat source, we consider a thermal capacity of the source being very large (an infinite heat source), and the thermal capacity of the sensor small. Therefore, the surface temperature T_b of a target can be considered constant during the measurement, while temperature of the sensor T_s is a function of time. That time function is dependent on the sensing element: its density, specific heat and thickness as per (3.82). If the input thermal flux has shape of a step function of time and the sensor is freely mounted in air, the output current can be approximated by an exponential function, so that $i = i_0 e^{-t/\tau}$ (8)

where i_0 is the peak current. In Fig. 3.29, as long as heat source is present, the charge Q and voltage V do not completely return to zero, no matter how much time has elapsed. Thermal energy enters the pyroelectric material from side b (Fig. 3.26), resulting in a material temperature increase. This causes the sensor's response, which decays with a thermal time constant τ_T. However, since the other side a of the sensor faces a cooler environment, part of the thermal energy leaves the sensor and is lost to its surroundings. Because the sides a and b face objects of different temperatures (one is a temperature of a source and the other is a temperature of the environment), a continuous heat flow exists through the pyroelectric material. Electric current generated by the pyroelectric sensor has the same shape as the thermal current through its material. An accurate measurement can demonstrate that as long as the heat continues to flow, the pyroelectric sensor will generate a constant voltage v_0 whose magnitude is proportional to the heat flow, thus making the device a heat flow sensor.

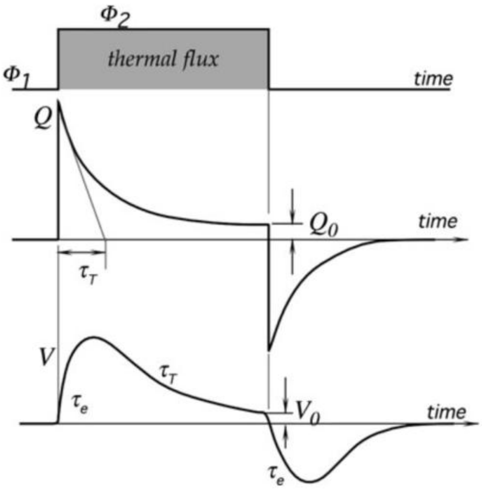


Fig. 3.29 Response of a pyroelectric sensor to a thermal step function The magnitudes of charge Q_0 and voltage V_0 are exaggerated for clarity

4.12. Semiconductor based Infrared Sensor

An IR sensor can sense its ambience using IR radiation usually emitted by the body itself to be detected. Heat or IR radiation are continuously radiated by objects at all temperatures, hence their position and motion can be detected and measured by IR sensors. Thermal detectors are much less sensitive in the mid-IR range, whereas quantum detectors work efficiently in this region. The operating principle of IR sensors is carried out by a step-by-step conversion of thermal radiation into heat, heat level or heat flow into an electrical signal.

An IR temperature sensor in noncontact mode has:

1. A sensing component, which is sensitive and receptive to the IR wavelength range of the electromagnetic spectrum. Thus the component must have rapid strong response to thermal radiation, and a high-quality long-standing stability.

2. An arrangement to support and carry the sensing component and provision to illuminate it to the radiation and small value of thermal conductivity to reduce the heat loss.

3. A suitable IR LED or source of light of appropriate wavelength which can be made incident.

4.12.1 Types of infrared sensors

IR sensors can be categorized into active and passive sensors: Thermal infrared sensors—IR radiation emitted by objects, also called heat, is detected and photosensitivity is not dependent on the wavelength being detected. Thermal detectors can be operated at room temperature or the object temperature, however, they have sluggish response times and lower detection abilities. The most

widely used thermal IR sensors are:

1. Golay cell.

2. Thermopile.

3. Pyroelectric detectors.

4. Bolometers.

5. Active FIR sensors.

In a nutshell, thermal energies possessed by objects radiate as IR radiation and behave as an IR source above absolute zero. Blackbody radiators, tungsten lamps etc, are sources of IR radiation, however, in most applications IR lasers and LEDs of specific wavelengths are used as IR sources in IR sensors.

Similar to all optical sensors, an emitter circuit and a receiver circuit are an ntegral part of an IR sensor. Thus, this combination is recognized as a photocoupler or an opto-coupler. An IR LED is the photo emitter, and an IR photodiode is the detector in this arrangement. The IR photodiode must have an appreciable responsivity for all IR radiations be it from an LED or any object. Net resistance of the circuit and thus the resultant output voltage of a photodiode responds linearly to the IR radiation reflection from a surface.

IR sensors too work on either direct or indirect incidence. The source of IR, an LED, is positioned in line to an LDR with no hindrance in the path, or is placed alongside with a reflecting surface opposite to it. The radiation transmitted strikes the opaque surface and reflects back to the photodiode, as shown in figure 3.19.

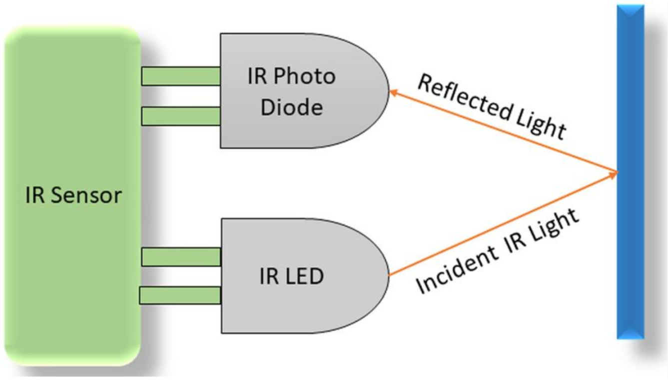


Figure 3.19. Schematic of a retro-reflective IR sensor.

Applications:

It is used as

proximity sensor
Line follower robots
Item counter
Burglar alarm
IR music transmitter and receiver
Night vision devices
Infrared astronomy