

**UNIT-4**  
**(SEMICONDUCTOR PHYSICS AND COMPUTATIONAL METHODS)**

## **Concepts of Electrical measurements & Two point probe technique**

### **General Introduction:**

Electrical property of the materials is one of the most important properties, which helps to classify the materials. For instance, solids may be classified in terms of their resistivity or conductivity as conductors, semiconductors or insulators.

**Ohm's law:** If physical conditions such as temperature, stress, strain remains unchanged in the materials, then potential difference ( $V$ ) across two ends of a conductor is proportional to current ( $I$ ) flowing through a conductor, i.e.,

$$V \propto I \text{ (or)}$$

$$V = I R$$

The constant of proportionality,  $R$ , is called resistance of the material.

#### **Resistivity:**

At a given constant temperature, the resistance  $R$  of the conductor is (i) proportional to its length ( $L$ ) and (ii) inversely proportional to its area of cross-section ( $A$ ), i.e.

$$R \propto L / A \text{ (or)} \quad R = \rho L / A$$

Resistivity is important for devices because it contributes to the device series resistance, capacitance, threshold voltage, hot carrier degradation of MOS devices, latch up of CMOS circuits, and other parameters.

- A material's conductivity,  $\sigma$ , (or the inverse property, resistivity,  $\rho = 1/\sigma$ ), relates to its ability to conduct electricity. In metals, conduction of electricity is tantamount to conduction of electrons, which depends on charge density and on scattering of the electrons by the crystal lattice (phonons) or by lattice imperfections.
- In semiconductors, conductivity is determined by the number of available charge carriers (electrons or holes) and the carrier mobility. Because of the different mechanisms for conductivity, its dependence on temperature also differs. Conductivity increases with increasing temperature for semiconductors (more carriers are generated) and it decreases with increasing temperature for metals (more scattering by the lattice).
- Conductivity also depends on physical structure. In crystals, the crystal type and orientation affect conductivity because the electronic structure is intimately tied to the crystal structure. The size of the crystallites (grains) in polycrystalline materials is also important as it affects the scattering of carriers, and, at very small sizes may also affect electronic structure.

**Table 1. General Characteristics for Conductivity Measurement Methods**

Method	Favored material Type	Favored Material Form	Estimated Measurement Range
Two-point measurement	High-resistance metals	Solid bar	$10^2 - 10^9 \Omega$
Four-point measurement	Metals	Solid bar	$10^{-7} - 10^4 \Omega$
Four-point probe	Semiconductor surface, thin metallic films	Planar solid	$10^{-3} - 10^5 \Omega\text{-cm}$
Van der Pauw	Semiconductor surface, thin metallic films	Planar solid	$10^{-3} - 10^5 \Omega\text{-cm}$

Resistivity is important for devices because it contributes to the device series resistance, capacitance, threshold voltage, hot carrier degradation of MOS devices, latch up of CMOS circuits, and other parameters. The wafers resistivity is usually modified locally during device processing by diffusion and ion implantation.

Two general methods involved for measuring the resistance of the materials are:

- (i) Two point probe method
- (ii) Four point probe method: **linear method & Van der pauw method**

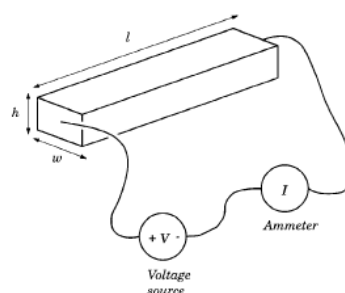
## Two-probe method:

Let us consider a rectangular bar of length  $l$ , height  $h$  and width  $w$  as shown in figure. copper wire are attached both ends of the bar.

The resistivity of the bar can be measured by measuring voltage drop across the wire due to passage of known current supplied by the battery  $E$  through the probes 1 and 2. The potential difference ( $V$ ) between the two contacts at the ends of the bar can be measured by a voltmeter. Therefore, the resistivity of the wire is, i.e.

$$\rho \equiv \frac{Rwh}{l}$$

In general, we use a multimeter for measuring the resistance of the materials. The typical range of resistance measured using the multimeter is  $1 \Omega$  to  $2 \text{ M}\Omega$ , but varies with the models and company.



While the two-probe method is a simple and advantageous method for measuring resistance above  $1\Omega$  directly, this method suffers from certain issues.

#### Two-probe method:

- Error due to contact resistance of the measuring leads,
- Materials having random shapes,
- Soldering of the test leads on some materials would be difficult,
- Heating of the leads during soldering may inject additional impurities in materials such as semiconductors and thereby affecting the intrinsic electrical resistivity largely.

In order to overcome the above problems, four-probe method is widely proposed.

### Four-point probe technique-linear method & Van der Pauw method

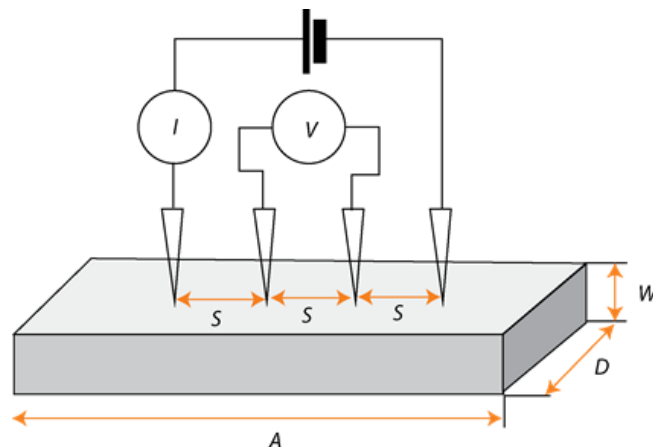
#### Four-probe method:

This method provides the measurement of the resistivity of the specimen having the wide variety of shapes but with uniform cross-section.

The soldering contacts proposed in two-probe method are replaced by pressure contacts to eliminate the last two problems summarized above.

In this method, four probes are utilized to measure the resistance of the samples. For example, two of the outer probes are used to send the current from the source meter and other two inner probes are used to measure the voltage drop across the sample.

The typical set up of the four-probe method is shown in Figure. There are four equally spaced tungsten metal tips supported by springs at one end to mount the sample surface without any damage.

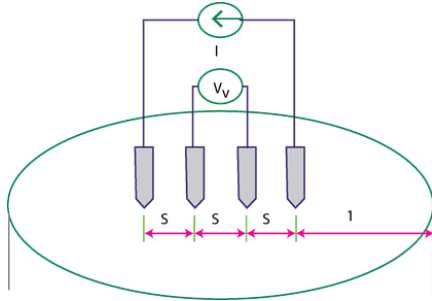


Two common techniques used in four-probe method are (i) Four-point collinear probe method and (ii) van der Pauw method.

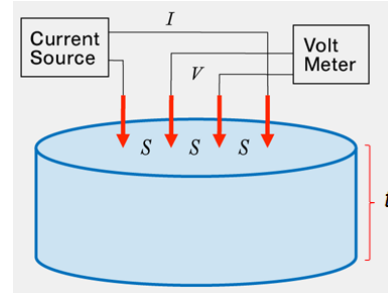
A high impedance current source is used to supply current through the outer two probes, which sets up an electric field in the sample. The potential difference developed across the inner probes, which draw no current due to the high input impedance voltmeter in the circuit, is measured through two inner probes.

### Four-point collinear probe method:

This is the most common way of measuring the resistivity of a material, which involves four equally spaced probes as shown in Figure, in contact with a materials of unknown resistance. This method can be used either in bulk or thin film specimen.



Schematic of four-point collinear probe method.



Schematic of four-point collinear probe method on bulk material.

### Four-point collinear probe method: For Bulk:

Consider a bulk material as shown in Figure , where the thickness ( $t$ ) of the materials is much higher than the space between the probes ( $s$ ), then the differential resistance due to spherical protrusion of current emanating from the outer probe tips is

$$\Delta R = \rho \left( \frac{dx}{A} \right)$$

Carrying out the integration between the inner probe tips,

$$R = \int_{v1}^{v2} \rho \frac{dx}{2\pi x^2} = \int_s^{2s} \rho \frac{dx}{2\pi x^2}$$

$$R = \frac{\rho}{2\pi} \left( -\frac{1}{x} \right) \Big|_s^{2s} = \frac{\rho}{2\pi} \frac{1}{2s}$$

where probe spacing is uniform. Due to the superposition of current at outer tips,  $R = V / (2I)$ . Therefore,

$$\rho = \left( \frac{V}{I} \right) (2\pi s)$$

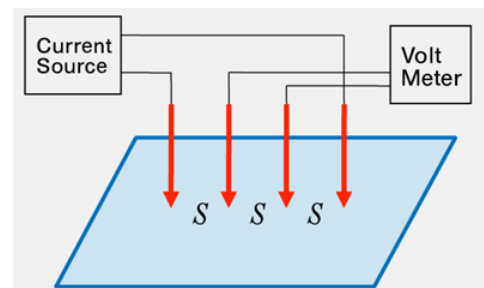
### Four-probe method: For Thin sheet:

For a very thin layer as shown in Figure, where the thickness of the sheet  $t \ll$  the space between the probes,  $s$ , we can get current rings instead of spheres. Therefore, the expression for the area is  $A = 2\pi x.t$ . Therefore, the derivation for resistance turns out to be:

$$R = \int_{v1}^{v2} \rho \frac{dx}{2\pi x t} = \int_s^{2s} \frac{\rho}{2\pi t} \frac{dx}{x}$$

$$R = \frac{\rho}{2\pi t} \ln(x) \Big|_s^{2s} = \frac{\rho}{2\pi t} \ln 2$$

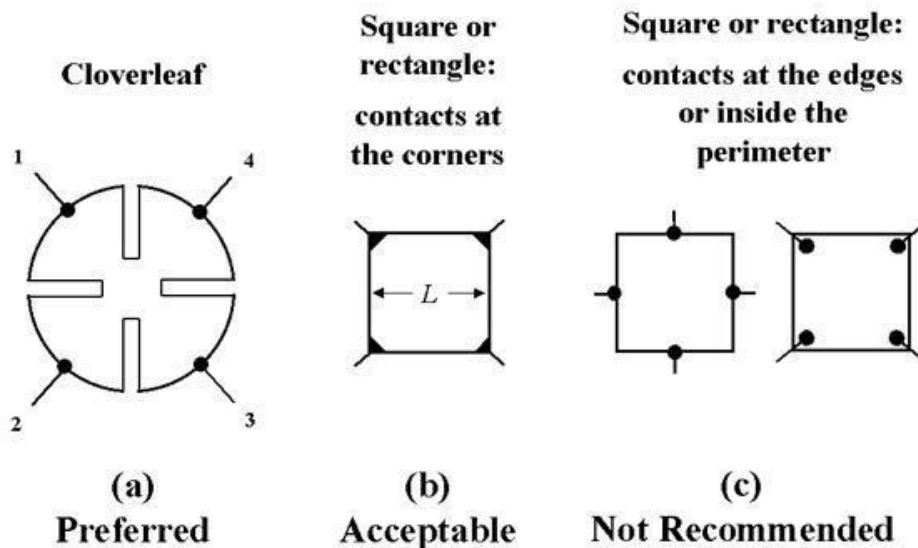
where probe spacing is uniform. Due to the superposition of current at outer tips,  $R = V / (2I)$ . Therefore,



$$\rho = \frac{Vt\pi}{I \ln 2}$$

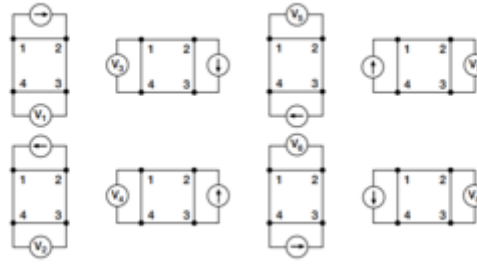
## Vander Pauw method

- The van der Pauw method is a technique commonly used to measure the resistivity and the Hall coefficient of a sample. There are five conditions that must be satisfied to use this technique:
- The sample must have a flat shape of uniform thickness. The sample must not have any isolated holes. The sample must be homogeneous and isotropic.
- All four contacts must be located at the edges of the sample. The area of contact of any individual contact should be at least an order of magnitude smaller than the area of the entire sample.
- In order to use the van der Pauw method, the sample thickness must be much less than the width and length of the sample. In order to reduce errors in the calculations, it is preferable that the sample be symmetrical. There must also be no isolated holes within the sample.



Make four ohmic contact with 8 periphery. The contacts are numbered from 1 to 4 in a counter-clockwise order, beginning at the top-left contact.

The resistivity can be derived from a total of eight measurements that are made around the periphery of the sample with the configurations shown in Figure.



Once all the voltage measurements are taken, two values of resistivity,  $\rho_A$  and  $\rho_B$ , are derived as follows:

$$\rho_A = \frac{\pi}{\ln 2} f_A t_s \frac{(V_1 - V_2 + V_3 - V_4)}{4I}$$

$$\rho_B = \frac{\pi}{\ln 2} f_B t_s \frac{(V_5 - V_6 + V_7 - V_8)}{4I}$$

where:  $\rho_A$  and  $\rho_B$  are volume resistivity in ohm-cm;  $t_s$  is the sample thickness in cm;  $V_1$ – $V_8$  represents the voltages measured by the voltmeter;  $I$  is the current through the sample in amperes;  $f_A$  and  $f_B$  are geometrical factors based on sample symmetry. They are related to the two resistance ratios  $Q_A$  and  $Q_B$  as shown in the following equations ( $f_A = f_B = 1$  for perfect symmetry).

Once  $\rho_A$  and  $\rho_B$  are known, the average resistivity ( $\rho_{AVG}$ ) can be determined as follows:

$$\rho_{AVG} = (\rho_A + \rho_B)/2$$

Following properties of a material can be measured by this method:

1. Resistivity
2. doping type (p- or n-type)
3. sheet density of majority carriers (majority carriers per unit area)
4. charge density and doping level.
5. mobility of majority carriers.

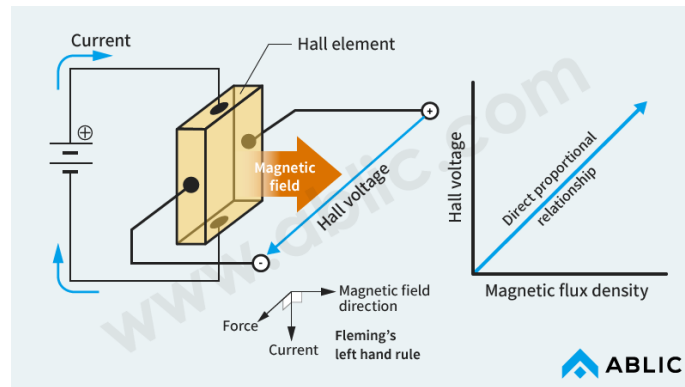
#### **Advantage of four probe method over two probe method**

Four point probe is preferred than two-point probe as the contact and spreading resistances in two point probe are large and the true resistivity cannot be actually separated from measured resistivity. In the four probe method, contact and spreading resistances are very low with voltage probes and hence accuracy in measurement is usually very high. To measure very low resistance values, four probe method is used. The resistance of probe will be not be added to that of sample being tested. It uses two wires to inject current in the resistance and another two wires to measure the drop against the resistance.

# HALL EFFECT

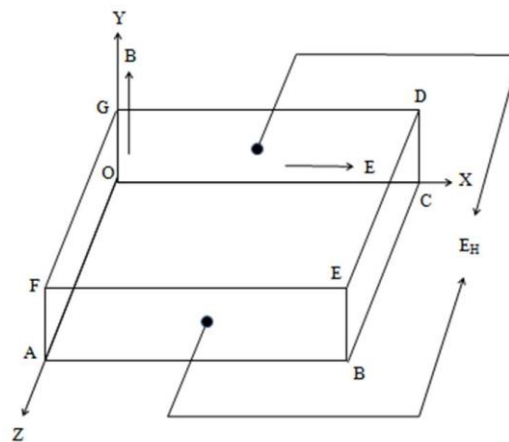
## Definition

When a piece of conductor (metal or Semiconductor) carrying current is placed in a transverse magnetic field, an electric field is produced inside the conductor in a direction normal to both the current and the magnetic field. This phenomenon is known as the Hall Effect and the generated voltage is called the Hall voltage.



## Explanation:

Consider a conventional current flow through the strip along OX and a magnetic field if induction  $B$  is applied along axis OY.



### Case – I If the Material is N-type Semi- Conductor (or) Metal

- If the strip is made up of metal or N-type semiconductor, the charge carriers in the strip will be electrons.
- As conventional current flows along OX, the electrons must be moving along XO. If the velocity of the electrons is ' $v$ ' and charge of the electrons is ' $-e$ ', the force on the electrons due to the magnetic field is,
- $F = -Bev$ , which acts along OZ. This causes the electrons to be deflected and so the electrons accumulate at the face ABEF.

- Thus, face ABEF will become negative and the face OCDG becomes positive. A potential difference is therefore established across faces ABEF and OCDG., causing a field  $E_H$ .
- This field gives rise to a force of ‘ $-eE_H$ ’ on the electrons in the opposite direction (i.e, in the negative Z direction).

At equilibrium ,  $eE_H = Bev$  (or)  $E_H = Bv$  (1)

If J is the current density, then ,  $J = -nev$  (2)

Where ‘n ‘ is the concentration of current carriers,

From equ. (2)

$v = J / -ne$  (3)

Substituting the value of v in equ. (1) we get,

$E_H = BJ / -ne$  (4) .

The Hall Effect is described by means of the Hall coefficient ‘ $R_H$  ‘ in terms of current density ‘ J ’ by the relation,

$E_H = R_H BJ$

(or)  $R_H = E_H / BJ$  (5)

By substituting the value of  $E_H$  from equ. (4) we

get,

$R_H = BJ / -neBJ = - 1/ne$  (6)

Since all the three quantities  $E_H$ , J and B are measurable , the Hall coefficient  $R_H$  and hence the carrier density ‘n’ can be found out.



### Case – (ii) If the material is a P-type semiconductor

If the strip is a P- type semiconductor, the charge carriers in the strip will be holes i.e., positively charged particles. The holes will constitute current in the direction of conventional current. Therefore, holes move along the direction of the conventional current itself i.e., along OX. If 'e' is the charge of the hole, the force experienced by the holes due to magnetic field is ,  $F = Bev$ , which acts along OZ. This causes the holes to accumulate on the face ABEF- making it positive and leaving the face OCDG as negative.

$$\text{Therefore , for a P-type semiconductor, } R_H = 1/pe \quad (7)$$

Where  $p$ = the density of holes.

### Determination of Hall Coefficient

The Hall coefficient is determined by measuring the Hall voltage that generates the Hall field. If 'w' is the width of the sample across which the Hall voltage is measured, then

$$E_H = V_H / w \quad (8)$$

$$\text{We know that } R_H = E_H / BJ \quad (9)$$

Substituting the value of  $E_H$  in the above equation, we get,

$$R_H = V_H / wBJ$$

(or)

$$V_H = R_H w BJ \quad (10)$$

If the thickness of the sample is 't' , then its cross sectional area  $A = wt$ , and the current density,

$$J = I/A = I/wt \quad (11)$$

Substitute the value of ' J ' in equation (11) , we get

$$V_H = R_H w BJ / wt = R_H I B / t$$

$$\text{(or) } R_H = V_H t / IB \quad (12)$$

$V_H$  will be opposite in sign for P and N type semiconductors

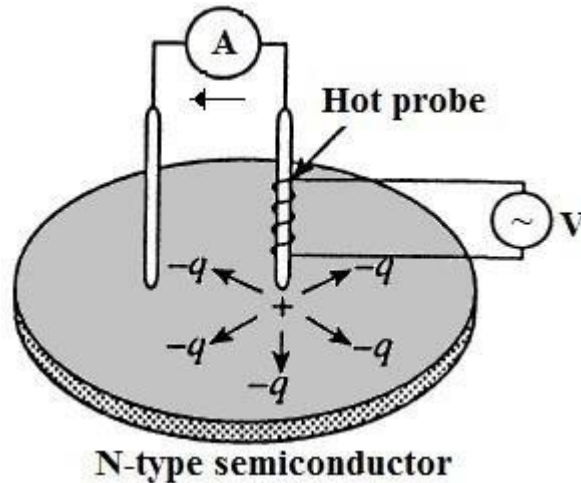
# HOT PROBE METHOD FOR SEMICONDUCTOR THIN FILM

- Physical properties of thin films significantly differ from those of bulk material
- There are various parameters such as a thickness, crystal structure, composition and other, which characterize a semiconductor film
- The parameters of charge carriers are
  - (a) Type of semiconductor,
  - (b) Impurities concentration,
  - (c) Mobility of charged carriers,
  - d. Diffusion coefficient.

Define the possibility to apply material for various electronic devices

## Principle:

- A conventional Hot-Probe experiment enables a simple and efficient way to distinguish between n-type and p-type semiconductors using a hot probe and a standard multi- meter.
- While applying the cold and hot probes to an n-type semiconductor, positive voltage readout is obtained in the meter, whereas for a p- type semiconductor, negative voltage is obtained.



## Experiment:

- A couple of a cold probe and a hot probe are attached to the semiconductor film surface.
- The hot probe is connected to the positive terminal of the multimeter while the cold probe is connected to the negative terminal.
- The thermally excited majority free charged carriers are translated within the semiconductor from the hot probe to the cold probe.

- Mechanism for this motion within the semiconductor is of a diffusion type since the material is uniformly doped due to the constant heating in the hot probe contact.
  - These translated majority carriers define the electrical potential sign of the measured current in the multimeter
- The Hot-Probe measurement may be described as a three-step process:
    1. The heated probe excites additional free charged carriers of two types
    2. The hot majority carriers begin to leave the heated part of the semiconductor surface by a diffusion mechanism. Simultaneously, a built-in electrical field is created between the electrodes and the second (cold) electrode is warmed as well. This warming and the built-in electrical field tend to prevent the diffusion process up to a halt at a steady state. This steady state condition exists until the heated source is switched off.
    3. The third process is actually a recombination of the excited additional charged carriers.

• This three-step process may be described, in general, by the continuity and Poisson's equations

$$\nabla J + \frac{\partial Q}{\partial t} = 0$$

$$\nabla E = \frac{Q}{\epsilon_0 \epsilon_r}$$

• Here Q is the uncompensated charge density excited by the heated electrode, J is the current density,  $\epsilon_0$  and  $\epsilon_r$  are the absolute and relative permittivity, and E is the built-in electrical field.

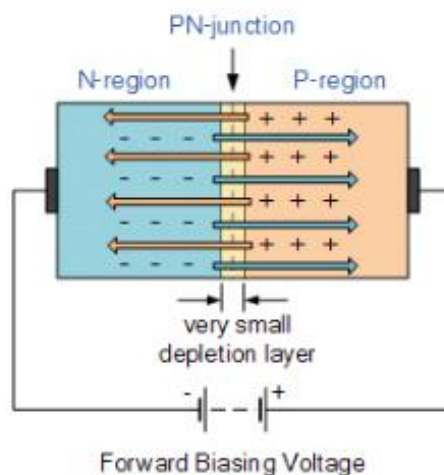
# Extraction Parameters & I-V Characteristic of PN Junction Diode

## PN Junction Diode

- If we join P-type and N-type semiconductor materials together they behave in a very different way and producing what is generally known as a “**PN Junction**”.
- The charge transfer of electrons and holes across the PN junction is known as **diffusion**. The width of these P and N layers depends on how heavily each side is doped with acceptor density  $N_A$ , and donor density  $N_D$ , respectively.
- Due to diffusion processes **potential barrier** is formed around the area of the junction as the donor atoms repel the holes and the acceptor atoms repel the electrons.
- At room temperature the voltage across the depletion layer for silicon is about 0.6 – 0.7 volts and for germanium is about 0.3 – 0.35 volts. This potential barrier will always exist even if the device is not connected to any external power source, as seen in diodes.

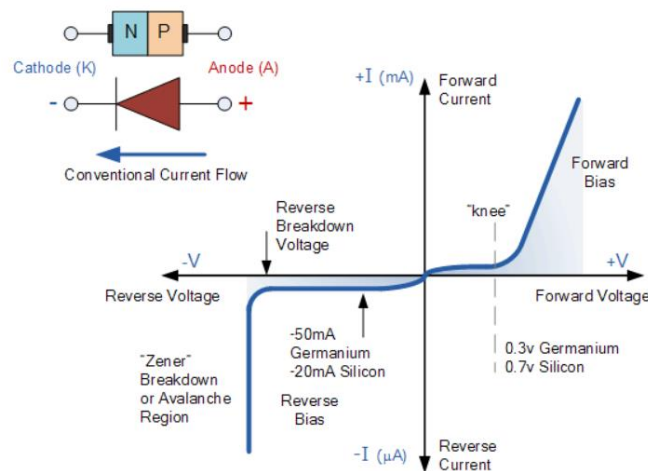
## I-V Characteristics in a Diode

- A *PN Junction Diode* is one of the simplest semiconductor devices around, and which has the characteristic of passing current in only one direction only. However, unlike a resistor, a diode does not behave linearly with respect to the applied voltage as the diode has an exponential current-voltage (I-V) relationship and therefore we can not described its operation by simply using an equation such as Ohm's law.
- If a suitable positive voltage (forward bias) is applied between the two ends of the PN junction, it can supply free electrons and holes with the extra energy they require to cross the junction as the width of the depletion layer around the PN junction is decreased.
- By applying a negative voltage (reverse bias) results in the free charges being pulled away from the junction resulting in the depletion layer width being increased. This has the effect of increasing or decreasing the effective resistance of the junction itself allowing or blocking current flow through the diode.

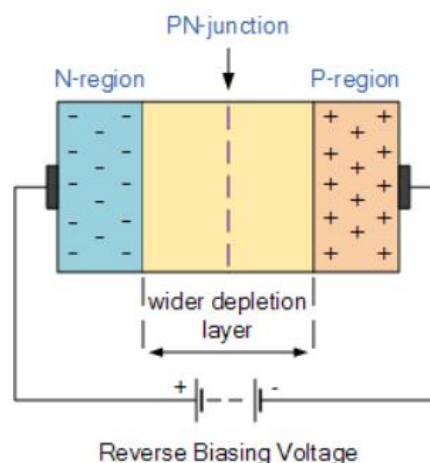


- When a diode is connected in a **Forward Bias** condition, a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material. If this external voltage becomes greater than the value of the potential barrier, approx. 0.7 volts for silicon

and 0.3 volts for germanium, the potential barriers opposition will be overcome and current will start to flow.



- Due to forward biasing the negative voltage pushes or repels electrons towards the junction giving them the energy to cross over and combine with the holes being pushed in the opposite direction towards the junction by the positive voltage. This results in a characteristics curve of zero current flowing up to this voltage point, called the “**knee**” on the static curves and then a high current flow through the diode with little increase in the external voltage.
- The point at which this sudden increase in current takes place is represented on the static I-V characteristics curve above as the “knee” point.
- Since the diode can conduct “infinite” current above this knee point as it effectively becomes a short circuit, therefore resistors are used in series with the diode to limit its current flow. Exceeding its maximum forward current specification causes the device to dissipate more power in the form of heat than it was designed for resulting in a very quick failure of the device.



- When a diode is connected in a **Reverse Bias** condition, a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material.
- The net result is that the depletion layer grows wider due to a lack of electrons and holes and presents a high impedance path, almost an insulator. The result is that a high

potential barrier is created thus preventing current from flowing through the semiconductor material.

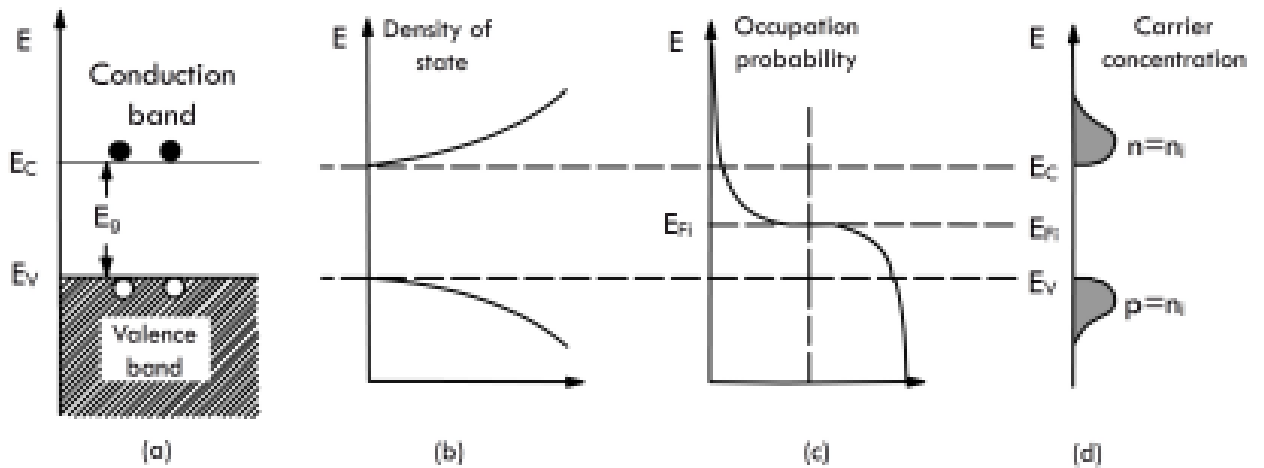
- This condition represents a high resistance value to the PN junction and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small **leakage current** does flow through the junction which can be measured in micro-amperes, ( $\mu\text{A}$ ).
- If the reverse bias voltage applied to the diode is increased to a sufficiently high enough value, it will cause the diode's PN junction to overheat and fail due to the avalanche effect around the junction. This may cause the diode to become shorted and will result in the flow of maximum circuit current, and this shown as a step downward slope in the reverse static characteristics curve below.
- When the reverse voltage is increased beyond the limit, then the reverse current increases drastically. This particular voltage that causes the drastic change in reverse current is called **Reverse Breakdown voltage**. Diode breakdown occurs by two mechanisms: Avalanche breakdown and Zener breakdown.

## Significance of band gap in semiconductors & Concept of absorption and emission

Opto electronic devices are based on the interaction of light with matter, typically semiconductors. In these devices, light or electromagnetic (EM) radiation is absorbed by the semiconductor and converted into electrical signals (electron-hole pairs) or electrical signals are converted to EM radiation (typically in the IR-Vis UV range). To understand the working of these devices it is essentially to look at light - semiconductor interaction. Consider light of wavelength,  $\lambda$ , incident on a semiconductor with band gap  $E_g$ . The energy of the radiation  $E$  is given by  $hc/\lambda$ .

- **Absorption in semiconductors:**  
Within these energy level systems we can have a variety of mechanisms by which electrons (and holes) absorb optical energy. Most of these processes can occur in quantum wells, wires, and dots, as well as in bulk material.
- **Band-to-band:** an electron in the valence band absorbs a photon with enough energy to be excited to the conduction band, leaving a hole behind.
- **Band-to-exciton:** an electron in the valence band absorbs almost enough energy to be excited to the conduction band.
- The electron and hole it leaves behind remain electrically "bound" together, much like the electron and proton of a hydrogen atom.
- **Band-to-impurity or impurity to band:** an electron absorbs a photon that excites it from the valence band to an empty impurity atom, or from an occupied impurity atom to the conduction band.
- **Free carrier:** an electron in the conduction band, or hole in the valence band, absorbs a photon and is excited to a higher energy level within the same set of bands (i.e, conduction or valence).
- In quantum structures there can be photon absorption due to carriers being excited between the quantum levels within the same band (termed "intra-band"), as well as between the various quantum levels in one band and those in another "(inter-band)":

- Intra-band: these transitions can occur only between even and odd index levels and are only operative for light polarized parallel to the direction of quantization.
- That is, in a quantum well the light must be polarized normal to the well itself, and in the direction of the composition variation.
- Inter-band: inter-band transitions can occur between conduction and valence bands, or between different valence bands (light-hole, heavy-hole, and spin-off).
- There are transitions can be active for either polarization of the light, depending on the symmetries of the respective bands.



**Figure 1:** (a) Energy band diagram for a semiconductor showing the valence and conduction band edges.  
(b) The density of states monotonically increases from the band edges  
(c) Occupation probability of electrons and holes in the CB and VB, calculated with respect to the intrinsic Fermi function  
(d) Product of (b) and (c) gives the concentration of electrons in CB and holes in the VB.

# Capacitance-Voltage measurements

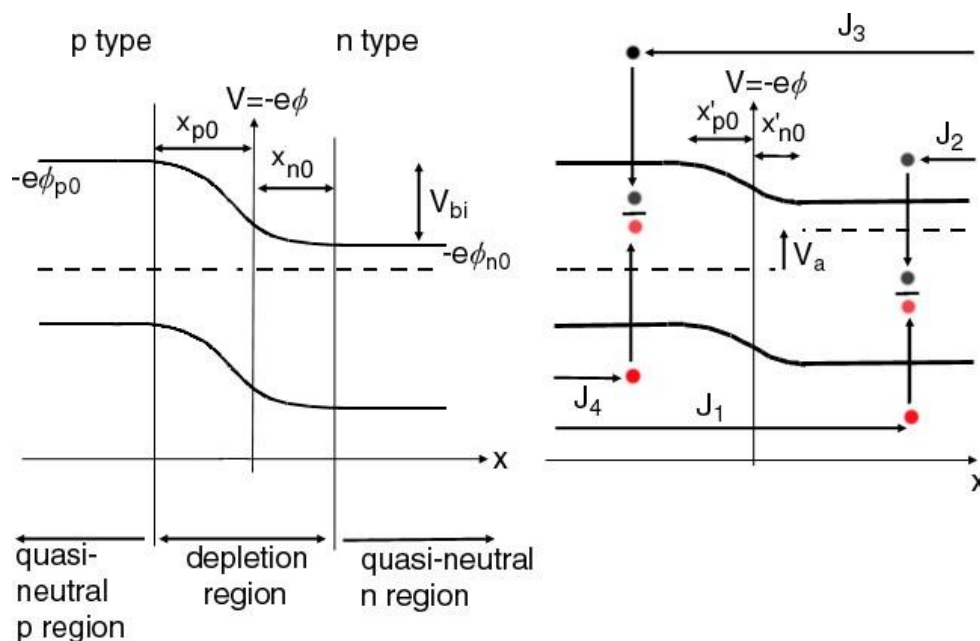
- Hillibrand and Gold (1960) first described the use of capacitance –voltage (C-V) methods to determine the majority carrier concentration in semiconductors.
- C-V measurements are capable of yielding quantitative information about the diffusion potential and doping concentration in semiconductor materials.
- The technique employs PN-junctions, metal- semiconductor junctions (Schottky barriers), electrolyte –semiconductor junction MIS field effect semiconductors.
- C-V measurements yield accurate information about the doping concentrations of majority carriers as a function of distance (depth) from the junction.

## Principle:

- The capacitance at an p-n or metal –semiconductor junctions depends on the properties of the charge- depletion layer formed at the junction
- The depletion regions is the vicinity of the PN junction and is “depleted” of free carriers due to the drift field required to maintain charge neutrality.

## Experiment:

- As shown in figure an abrupt pn junction is considered.
- The bandgap of the semiconductor  $E_G = E_c - E_v$  is defined by the difference between the conduction band energy  $E_c$  and the valance band energy  $E_v$ .
- The fermi energy  $E_F$  defines the equilibrium condition for the charge neutrality.
- The difference in energy between the conduction band as one crosses the PN junction is called the diffusion potential,  $V_{bi}$  (built-in- potential).





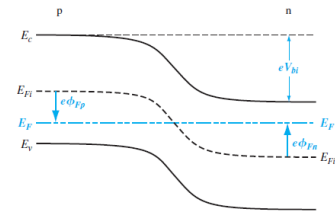
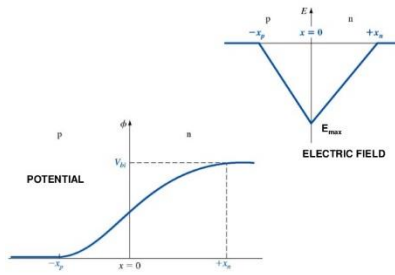


Figure 7.3 | Energy-band diagram of a pn junction in thermal equilibrium.

## Diagram Explanation

- Abrupt PN junction in thermal equilibrium (no bias).
- A. Space charge distribution in the depletion approximation.  
The dashed lines indicate the majority carrier distribution tails.
- B. Electric field across the depletion region
- C. Potential distribution due to the electric field where  $V_{bi}$  is the (built-in) diffusion potential
- D. Energy band diagram
- Consider the PN junction, where the regions denoted by + and – indicates the junction region depleted of free carriers, leaving behind ionized donors and acceptors.
- In this region from Poisson's equations

$$\bullet \quad \frac{\partial^2 v}{\partial x^2} = \frac{\partial E}{\partial x} = \frac{\rho(x)}{\epsilon} = \frac{e}{\epsilon} [p(x) - n(x) + N_D^+(x) - N_A^-(x)]$$

For predominantly doped p-type

$$\bullet \quad \frac{\partial^2 v}{\partial x^2} \approx \frac{e}{\epsilon} N_D^+ \quad \text{for } 0 < x \leq x_n$$

And for n-type

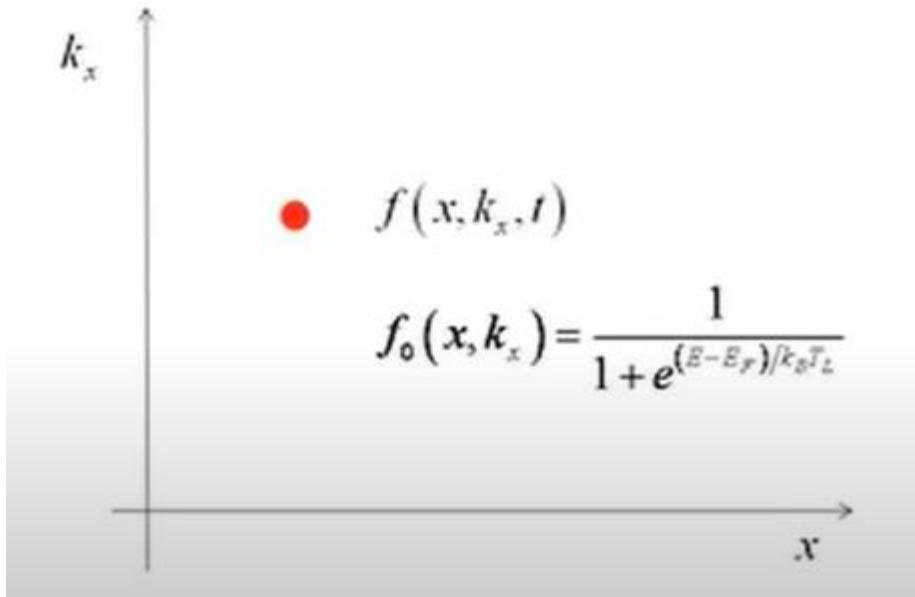
$$\bullet \quad \frac{\partial^2 v}{\partial x^2} \approx \frac{e}{\epsilon} N_A^- \quad \text{for } -x_p \leq x < 0$$

- Where  $v$  - Voltage,  $\epsilon$  - Electric field,  $e$  - charge,  $p(x)$  and  $n(x)$  - the hole and electron concentration (electric potential) comprising the mobile carriers.
- $N_D^+$  and  $N_A^-$  - the donor and acceptor doping concentrations
- $\epsilon = K_s \epsilon_0$  - The permittivity with dielectric coefficients
- The spatial dependence,  $x$  is measured relative to the physical location of the p-n junction. The solution of three equations is a form useful for C-V measurements is
- $V(x) = V_{bi} \left[ 2 \left( \frac{x}{W} \right) - \left( \frac{x}{W} \right)^2 \right]$
- where  $V_{bi} = \frac{KT}{e} \ln \left( \frac{N_d N_a}{n_i^2} \right)$

# Boltzmann Transport equation

**Boltzmann Transport equation (BTE)** is very useful in understanding the transport properties such as electrical conductivity, thermal conductivity, and thermoelectric power.

This equation is used to determine the distribution function of particles (electrons) in the phase space  $(\mathbf{r}, \mathbf{k})$ .



Consider simply a particle present in a simple 1D space (in phase space 1D has two coordinates  $x$  and  $p_x$ ). The probability of a particle occupying a particular state at equilibrium condition can be expressed through Fermi function.

Question is how this probability or distribution function looks like in 3D space and how it affects with collision or force of the electrons?

BTE is a semi-classical equation considering electron as semi-classical particle which obeys following modified Newtons equations.

$$\frac{d(\hbar \vec{k})}{dt} = -\nabla_r E_C(r) = -q \vec{\xi}(\vec{r})$$

$\hbar \vec{k}$  - momentum

$$\hbar \vec{k}(t) = \hbar \vec{k}(0) + \int_0^t -q \vec{\xi}(t') dt'$$

$E_C$  - conduction band bottom energy

$$\vec{v}_g(t) = \frac{1}{\hbar} \nabla_k E[\vec{k}(t)]$$

$\vec{\xi}(\vec{r})$  - Electric field

$$\vec{r}(t) = \vec{r}(0) + \int_0^t \vec{v}_g(t') dt'$$

Consider the electron present in an energy band described by Bloch function  $\psi_k(\vec{r}) = |\vec{k}\rangle = e^{ik \cdot \vec{r}} u_k(\vec{r})$

The number of electrons per unit volume whose wavevectors lie in the interval  $(\vec{k} - \vec{k} + d\vec{k})$  is

$$\frac{2}{(2\pi)^3} f(\vec{k}, \vec{r}) d\vec{k}$$

where the factor 2 is the spin weight. In equilibrium  $f(\vec{k}, \vec{r})$  becomes the Fermi-Dirac distribution function  $f_0(E)$ , but deviates from  $f_0(E)$  in the presence of the electric field, magnetic field, temperature gradient, and so on. There are two contributions this time dependence, (i) from the external force (the drift term) and (ii) the collisions (the collision term),

$$\frac{df}{dt} = \left( \frac{\partial f}{\partial t} \right)_{\text{drift}} + \left( \frac{\partial f}{\partial t} \right)_{\text{coll}}$$

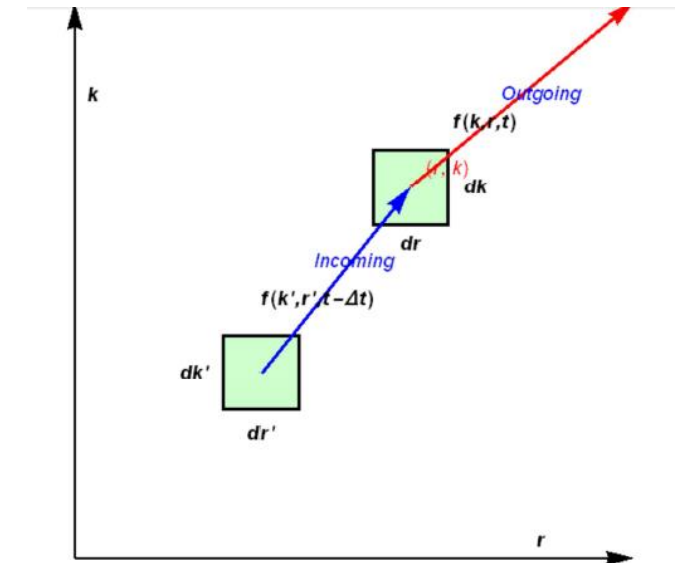
If we consider the evolution of a particle as in this graph, the function in the dimension  $dr dk$  is  $f(k, r, t)$  and the same function is  $f(k', r', t - \Delta t)$  in  $dr' dk'$ .

If there is no collision only external force is applied, then

$$f(\vec{k}, \vec{r}, t) = f(\vec{k}', \vec{r}', t - \Delta t)$$

$$\frac{\partial f}{\partial t} = 0$$

Its called as Collision less  
BTE – Ballistic transport



The above equation can be written as

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{drift} + \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial p_x} \frac{\partial p_x}{\partial t} = 0$$

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{drift} + \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial p_x} F_x = 0$$

For three dimensions,

$$\left( \frac{\partial f}{\partial t} \right)_{drift} + \vec{v} \cdot \nabla_r f + \vec{F}_e \cdot \nabla_p f = 0$$

$$\left( \frac{\partial f}{\partial t} \right)_{drift} = -\vec{v} \cdot \nabla_r f - \vec{F}_e \cdot \nabla_p f = 0$$

$$\begin{aligned}\vec{F}_e &= -q\vec{E} - q\vec{v} \times \vec{B} \\ \nabla_r f &= \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \\ \nabla_p f &= \frac{\partial f}{\partial p_x} \hat{p}_x + \frac{\partial f}{\partial p_y} \hat{p}_y + \frac{\partial f}{\partial p_z} \hat{p}_z \\ \vec{p} &= \hbar \vec{k}\end{aligned}$$

Assumptions considered in above collision less BTE

1. Semiclassical treatment of electrons in a crystal with  $E(k)$
2. Neglected generation and recombination
3. Neglected e-e correlations (mean field approximation)

Whereas, the collision term can be written as

$$\left( \frac{\partial f}{\partial t} \right)_{coll} = - \frac{f - f_0(k)}{\tau(k)},$$

Where  $f_0$  – equilibrium distribution function

$f$  – distribution function after a collision

$\tau$  – relaxation time

We consider the general Boltzmann equation

$$\frac{df}{dt} = \left( \frac{\partial f}{\partial t} \right)_{drift} + \left( \frac{\partial f}{\partial t} \right)_{colli}$$

In the steady state,  $\frac{df}{dt} = 0$

$$\left( \frac{\partial f}{\partial t} \right)_{drift} + \left( \frac{\partial f}{\partial t} \right)_{colli} = 0$$

$$f = f_0 - \tau \vec{v} \cdot \nabla_r f - \frac{1}{\hbar} \tau \vec{F} \cdot \nabla_k f$$

$$\begin{aligned} \vec{F}_e &= -q\vec{E} - q\vec{v} \times \vec{B} \\ \nabla_r f &= \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \\ \nabla_p f &= \frac{\partial f}{\partial p_x} \hat{p}_x + \frac{\partial f}{\partial p_y} \hat{p}_y + \frac{\partial f}{\partial p_z} \hat{p}_z \\ \vec{p} &= \hbar \vec{k} \end{aligned}$$

$$\begin{aligned} -\vec{v} \cdot \nabla_r f - \frac{1}{\hbar} \vec{F} \cdot \nabla_k f - \frac{(f - f_0(k))}{\tau} &= 0 \\ -\vec{v} \cdot \tau \cdot \nabla_r f - \frac{1}{\hbar} \vec{F} \cdot \tau \cdot \nabla_k f + f - f_0(k) &= 0 \\ f &= f_0(k) - \tau \vec{v} \cdot \nabla_r f - \frac{1}{\hbar} \vec{F} \cdot \tau \cdot \nabla_k f \end{aligned}$$

This equation is called the Boltzmann transport equation.

# Introduction of TCAD : (Technology Computer-Aided Design)

## Physical modeling: definitions

### *Physical Modeling*

Representation of the physical behavior of a system (device) by an abstract mathematical model which approximates this behavior. Such a model may either be a closed-form expression (analytical model), or, in general, a system of coupled (differential) equations to be solved numerically.

### *Analytical Modeling vs. Numerical Modeling*

**Analytical modeling** basically means the representation of a physical property or law in terms of approximate closed-form expressions using “lumped” parameters. It is also called “compact” modeling.

**Numerical modeling:** modeling of the device behavior through the numerical solution of the differential equations describing the device physics on a given geometrical domain.

In the literature, the word “**modeling**” usually implies analytical/compact modeling, while “**simulation**” is much used for numerical modeling

# Examples

- Analytical modeling

- $I_{DS}$ - $V_{DS}$  curve of a MOS transistor

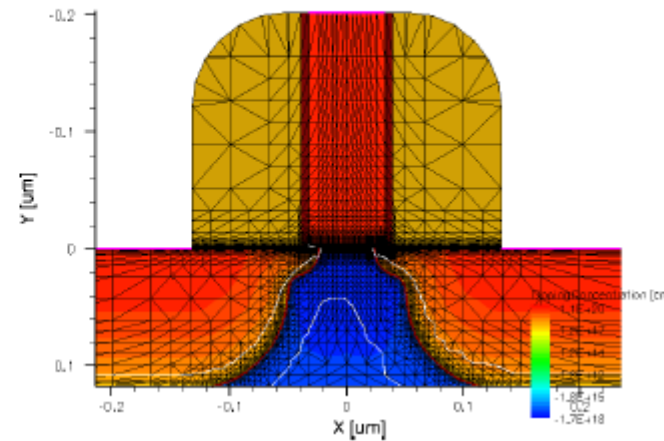
$$I_D = \mu_n C_{ox} \frac{W}{L} \left[ (V_{GS} - V_{TH}) V_{DS} - \frac{1}{2} V_{DS}^2 \right]$$

- Numerical modeling

- Drift-Diffusion numerical model

$$\left\{ \begin{array}{l} \nabla \cdot \mathcal{E} = \rho / K_S \epsilon_0 \\ \frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \mathbf{J}_N - r_N + g_N \\ \frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot \mathbf{J}_P - r_P + g_P \\ \mathbf{J}_N = q \mu_n n \mathcal{E} + q D_N \nabla n \\ \mathbf{J}_P = q \mu_p p \mathcal{E} - q D_P \nabla p \end{array} \right.$$

solved at each node of a discretized domain





# TCAD

Numerical simulation is much used to understand advanced device physics, for device design, scaling analyses & interaction with process manufacturing.

TCAD stands for “Technology Computer-Aided Design” is one of a numerical modelling method for semiconductor devices.

TCAD is a branch of Electronic Design Automation (EDA) that models semiconductor fabrication and semiconductor device operation. The modeling of the fabrication is termed *Process TCAD*, while the modeling of the device operation is termed *Device TCAD*. The aim of TCAD is the design of semiconductor processes and devices to fulfill some given specifications.

## Process TCAD:

**modeling of semiconductor-chip process-manufacturing steps** like lithography, deposition, etching, ion implantation, diffusion, oxidation, silicidation, mechanical stress, etc.

It requires detailed modeling of the *physical principles of manufacturing*, and usually also the modeling of the *specific equipments* used. Calibration of models needs expensive experiments (ad-hoc wafer fabrication, physical-chemical investigations).

## Device TCAD:

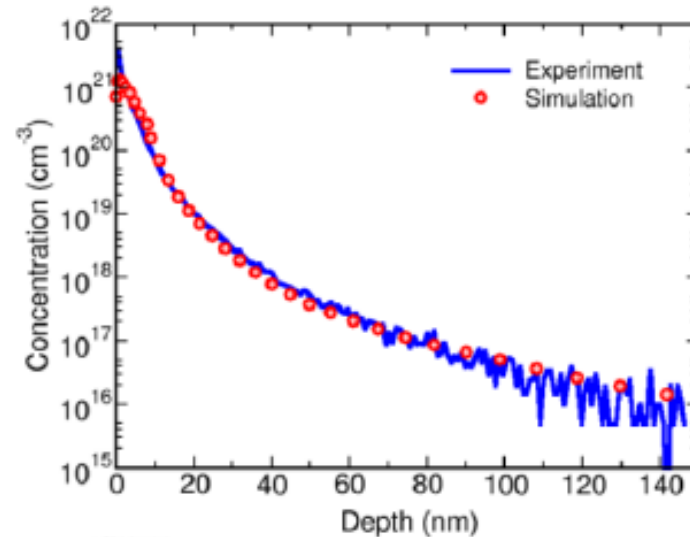
**modeling of electrical, thermal, optical and mechanical behavior of semiconductor devices** (e.g., diode, BJT, MOSFET, solar cell,...).

It focuses on the physical principles at the basis of *carrier transport* and of *optical generation* in semiconductor devices. Models are more easily generalized than for process physics. In addition, they do not need moving boundaries/moving meshes, as instead process simulations need, i.e. convergence is in general easier. Calibration of models usually needs only electrical characterization of fabricated samples.

# What is TCAD? – Examples

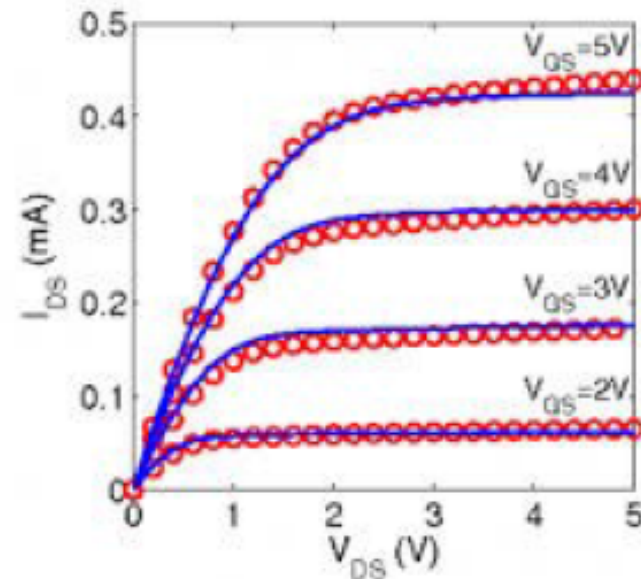
- Process simulations

Simulate doping profiles obtained by specific processing techniques, calibrate the model with experimental data and then optimize the process to obtain the desired profile.



- Device simulations

Simulate the output characteristics of a MOSFET device and calibrate the device architecture to fine-tune the device performance.



## Advantageous of TCAD

- The device features can be optimized if hands-on calculations are too complicated or impose unacceptable assumptions
- It helps to make predictions (scaling, new device concepts) when hands-on calculations are not viable (e.g., complex devices, modeling of distributed statistical effects or process yield).
- To get insights. No real experiment will probably be ever able to measure some of the physical quantities calculated by TCAD tools (e.g., local distribution of carriers, local electric field, etc.).
- To quickly screen technological options and drive the industrial strategy

# Monte Carlo method

## Introduction to MC method

- It have been invented in the context of the development of the atomic bomb in the 1940's.
- Monte carlo methods are a class of computational algorithms, it can be applied to vast ranges of problems.
- It is not a statistical tool, it rely on **repeated random sampling**.
- **It** provide generally approximate solutions, and are used in cases where analytical or numerical solutions don't exist or are too difficult to implement

MC methods have applications in various fields such as Material Science, Statistical Physics, Chemical and Bio Physics, nuclear physics, flow of traffic and many others.

*Monte-Carlo methods generally follow the following steps:*

1. Determine the **statistical properties** of possible inputs
2. Generate many **sets of possible inputs** which follows the above properties
3. Perform a **deterministic calculation** with these sets
4. Analyze **statistically** the results

The error on the results typically decreases as  $1 / \sqrt{N}$

## A simple example of MC

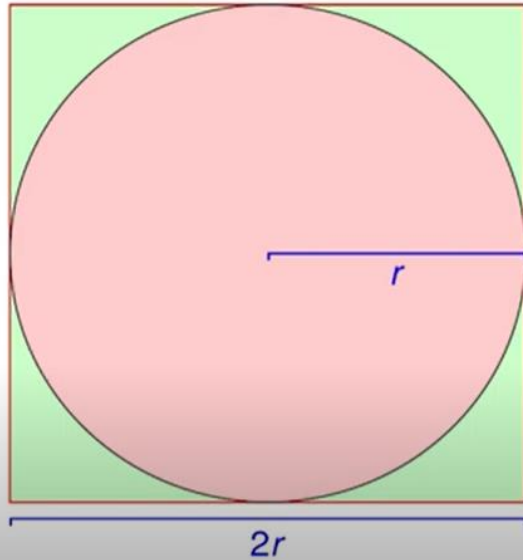
Can we determine the value of  $\pi$  using a Monte Carlo method ?

- Consider a square with sides of length  $2r$ .

$$S_{\text{square}} = (2r)^2 = 4r^2.$$

- Draw a circle inside the square with radius  $r$ .

$$S_{\text{circle}} = \pi r^2.$$



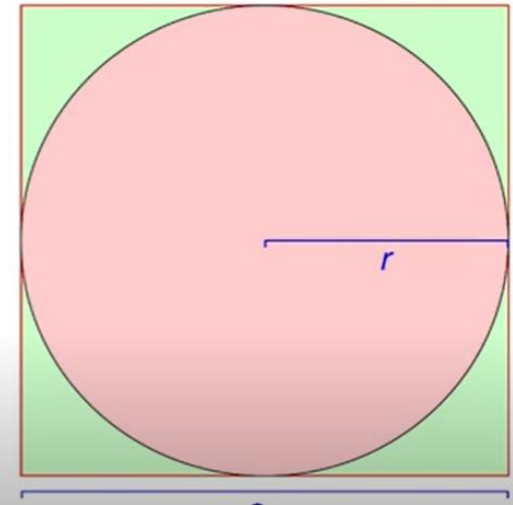
- Then the ratio is

$$\frac{S_{\text{circle}}}{S_{\text{square}}} = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}.$$

- Multiplying 4 on both sides of the equation yields

$$4 \times \frac{S_{\text{circle}}}{S_{\text{square}}} = 4 \times \frac{\pi}{4}$$

$$4 \times \frac{S_{\text{circle}}}{S_{\text{square}}} = \pi$$



$$\pi = 4 \times \frac{S_{\text{circle}}}{S_{\text{square}}}$$

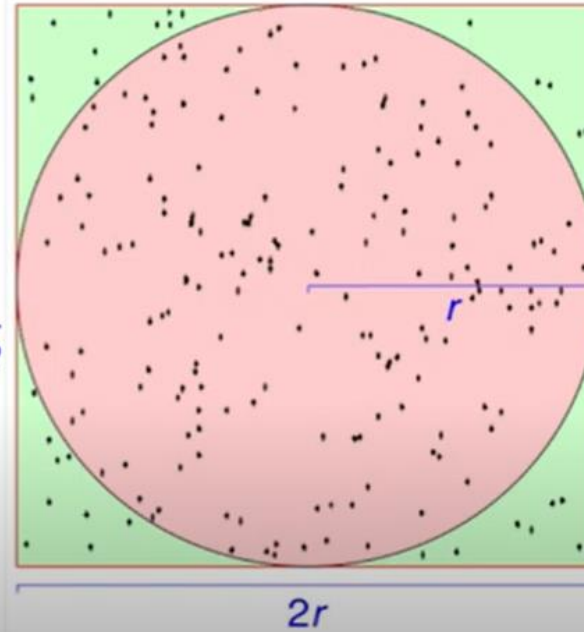
- Since  $\pi = 4 \times \frac{S_{\text{circle}}}{S_{\text{square}}}$

- We have

$$\pi = 4 \times \frac{S_{\text{circle}}}{S_{\text{square}}}$$

$\approx$

$$4 \times \frac{\text{Points within the circle}}{\text{Points within the Square}}$$



Draw a square on the ground, then inscribe a circle within it. Uniformly scatter some objects of uniform size (grains of rice or sand) over the square.

Count the number of objects inside the circle and the total number of objects.

The ratio of the two counts is an estimate of the ratio of the two areas, which is  $\pi/4$ . Multiply the result by 4 to estimate  $\pi$ .

## Solving Boltzmann Transport equation using Monte-Carlo method

Direct simulation Monte Carlo (DSMC – by Dr. Bird at 1963) method is one of a popular choice for solving Boltzmann transport equation (BTE) in the field of rarefied gas dynamics due to a number of factors:

First, the high dimensionality associated with  $f(t, x, p)$  makes numerical methods based on discretization computationally expensive, both in terms of CPU time and storage

second, the particle formulation employed by DSMC is ideal for accurately and *stably* capturing the propagation of traveling discontinuities in the distribution function resulting from the advection operator in the Boltzmann equation.

finally, the DSMC algorithm combines simplicity with an intuitive formulation that naturally employs importance sampling for improved computational efficiency



DSMC solves the Boltzmann equation using discretization in time; each time step of length  $\Delta t$  is split into a collisionless advection and a collision substep. Numerically, this corresponds to a splitting scheme in which the collisionless advection substep integrates

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{F} \cdot \nabla_{\mathbf{v}} f = 0$$

while the collision substep integrates

$$\frac{\partial f}{\partial t} = \left[ \frac{\partial f}{\partial t} \right]_{\text{coll}}$$

During the advection substep, particles move ballistically (according to their velocities and accelerations).

During the collision substep, the distribution function is updated by processing binary collisions between collision pairs chosen from within the same computational cell of linear size  $\Delta \mathbf{x}$ , using an acceptance-rejection procedure.

This introduces the second important mode of discretization in this algorithm, as it corresponds to treating the distribution function as spatially homogeneous within each cell.