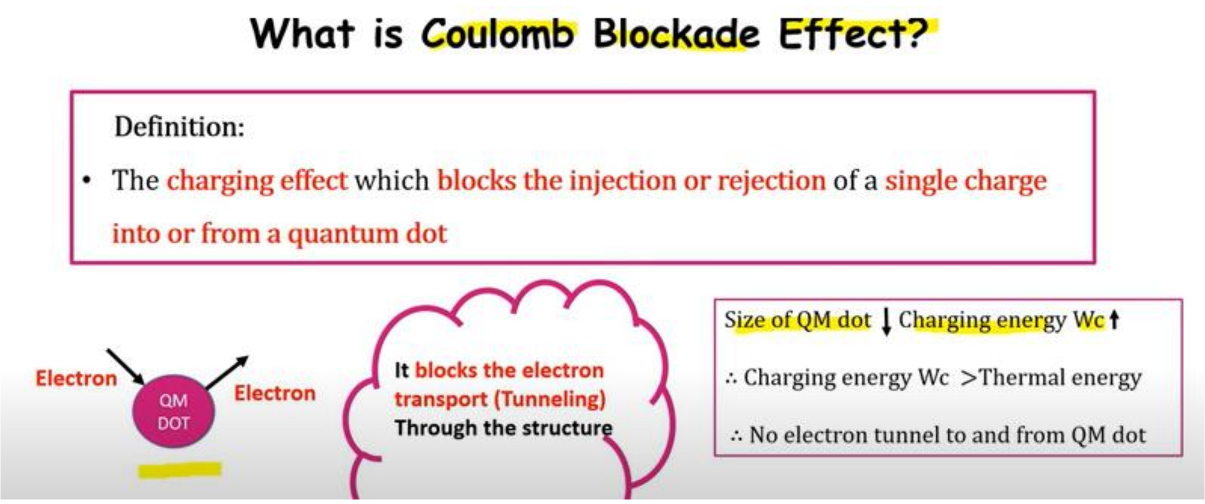


What is Coulomb Blockade Effect?

Definition:

- The charging effect which blocks the injection or rejection of a single charge into or from a quantum dot

It blocks the electron transport (Tunneling) Through the structure

Size of QM dot ↓ Charging energy Wc ↑

∴ Charging energy Wc >Thermal energy

∴ No electron tunnel to and from QM dot

Coulomb blockade

In mesoscopic physics, a **Coulomb blockade** (**CB**), named after Charles-Augustin de Coulomb's electrical force, is the decrease in electrical conductance at small bias voltages of a small electronic device comprising at least one low-capacitance tunnel junction. Because of the CB, the conductance of a device may not be constant at low bias voltages, but disappear for biases under a certain threshold, i.e. no current flows.

Coulomb blockade can be observed by making a device very small, like a quantum dot. When the device is small enough, electrons inside the device will create a strong Coulomb repulsion preventing other electrons to flow. Thus, the device will no longer follow Ohm's law and the current-voltage relation of the Coulomb blockade looks like a staircase.

Even though the Coulomb blockade can be used to demonstrate the quantization of the electric charge, it remains a classical effect and its main description does not require quantum mechanics. However, when few electrons are involved and an external static magnetic field is applied, Coulomb blockade provides the ground for a spin blockade (like Pauli spin blockade) and valley blockade, which include quantum mechanical effects due to spin and orbital interactions respectively between the electrons

A capacitor is a (macroscopic) device that stores electrical charge. The basic structure of a capacitor is shown at the right. When a voltage is applied to such a device, it develops a charge (± Q) on the two plates that is proportional to the voltage:

$$C=Q/V \quad (1)$$

The magnitude of the capacitance C is determined by the permittivity ε and the dimensions of the dielectric layer, A(area of parallel plate) and d(separation of plate).

$$C=\varepsilon A/d \quad (2)$$

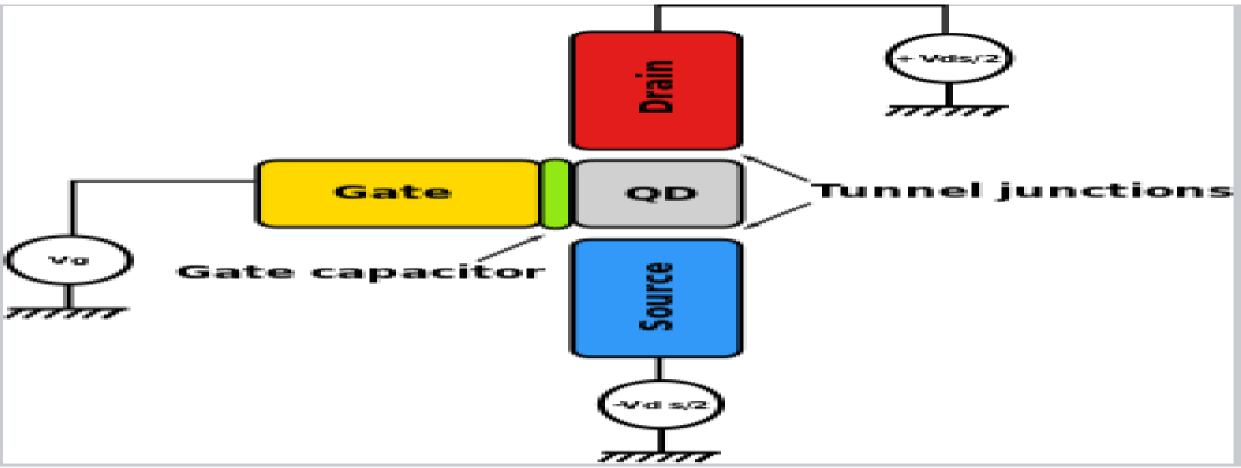
We can also calculate the work done in charging the capacitor up (i.e., the energy stored by charging the capacitor) by integrating the voltage times the charge:

$$W=\int_0^Q V\,dq = \int_0^Q \frac{q}{C}\,dq = Q^2/2C$$

Now it is interesting to ask, what happens to a capacitor when we make it very small? This is of particular interest in a device called a **single electron transistor**, a schematic diagram of which is shown fig 2. The metallic gate lead is separated from a "quantum dot," which can be a metal or semiconductor particle, by a thin dielectric layer. This metal-dielectric-metal sandwich acts as a capacitor, and from the equation above, the energy needed to charge it by a single electron (Q = e) is:

$$E = e^2/2C$$

where e is the charge of the electron, 1.602 x 10^{-19} Coulomb. If the gate width and lateral dimensions are very small - say 2 nm as is readily achievable in self-assembled Coulomb blockade devices - then for a typical insulating dielectric, a voltage of about 200 mV is needed to charge the quantum dot by a single electron. Again, this effect is unique to the nanoscale, because a 10 times larger device area would make the single-electron charging voltage about 20 mV, which is smaller than the thermal energy kT (26 meV). Thus for devices larger than about 5-6 nm, individual electron charging events are washed out at room temperature by thermal fluctuations.



Schematic of a single-electron transistor.

How can a **nanoscale capacitor** like this act as a **transistor**, which functions as a switch in an electrical circuit? The effect comes from the mutual repulsion of electrons. An electron on the quantum dot repels any other electron that would be forced onto it by applying a small voltage between the source and the drain. Hence the conductance of the quantum dot is very low at a gate bias of zero volts, or at any gate bias (200, 400, 600 mV...) that places an integer number (1, 2, 3,...) of electrons on the dot. But halfway in between these voltages (e.g., at 100, 300, 500 mV) the energy is the same whether there are n or n+1 electrons on the dot. This means that electrons can hop on and off without changing their energy, i.e., they can tunnel through the dot from source to drain. This effect gives peaks in the conductance of the dot at regular steps in the gate voltage. In effect, the gate can act as a switch, as in a conventional field-effect transistor. Single-electron transistors are being researched as ultra-sensitive electrometers and single-molecule chemical sensors, since a tiny change in the electrostatic environment of the dot can switch the device on or off.

Electrostatic energy E of a spherical capacitor of charge Q is changed by amount$\Delta \boldsymbol{E} = \frac{\boldsymbol{eQ}}{\boldsymbol{C}}$ when a single electron is added or subtracted. The corresponding change in potential $\Delta \boldsymbol{V} = \frac{\boldsymbol{e}}{\boldsymbol{C}}$

For a spherical capacitor of diameter d the capacitance C is given by

C= $2\pi\varepsilon d$ =$4\pi\epsilon$r

So $\Delta \boldsymbol{V} = \frac{\boldsymbol{e}}{\boldsymbol{C}} \propto \frac{\boldsymbol{1}}{\boldsymbol{r}}$

From the above equation as the **size of quantum dot is reduced** , capacitance C gets smaller, the value of energy necessary to change number of electrons in the dot increases.

Thus it is evident that for electron to enter the dot, the voltage will have to be raised to at least $\frac{\boldsymbol{e}}{\boldsymbol{2C}}$. So electron cannot tunnel if

|V| $\leq\leq \frac{\boldsymbol{e}}{\boldsymbol{2C}}$

Therefore , there is a voltage range, between **-e/2C** and **e/2C** in which current can not go through the quantum dot & hence the name **Coulomb Blockade has been given to the phenomena.**

Evidently if the above process is continued and we keep adding more electron, we will have the situation in which we will observe discontinuities in the current through the quantum dot whenever the voltage acquires the value expressed by

V = $\frac{\boldsymbol{1}}{\boldsymbol{2C}}(\boldsymbol{2n} + \boldsymbol{1})\boldsymbol{e}$ **, n=0,1,2,…….**

Electron can tunnel only at V_g for which the following condition is to be satisfied

E(N+1,V_g) = E(N,V_g) $\pm \boldsymbol{KT})$

Therefore the region of no current of low bias voltage is known as Coulomb Blockade region. Repeated tunnelling of single electrons produces what is known as Coulomb Staircase.

Coulomb blockade involves a small island of charge situation between two electrodes. If the island is small enough and has *N* electrons, then an energy gap opens up between the energy of the last (*N*th) electron and the first empty electron state (*N* + 1). This gap equals the square of the electron charge e^2, divided by the capacitance of the island *C*. Therefore if the island is small enough so that this energy gap is larger than the thermal energy in the system (k_BT where k_B is Boltzmann's constant and *T* is the temperature), then electrons cannot quantum mechanically tunnel through the system since the only free states that electrons may tunnel onto the island are above the energy of the electrons in the electrodes (Fig. 4). If, however, a gate is used to electrostatically move the islands energy states with respect to the electrodes then the $N + 1$th electron free state can be moved below that of the left electrode and electrons can quantum mechanically tunnel through the island one at a time [Fig. (4b)], hence the name single electron transistor. Therefore the current-voltage characteristics have zero current until the applied gate voltage corresponds to $\pm e / 2C$ [Fig. (4c)].

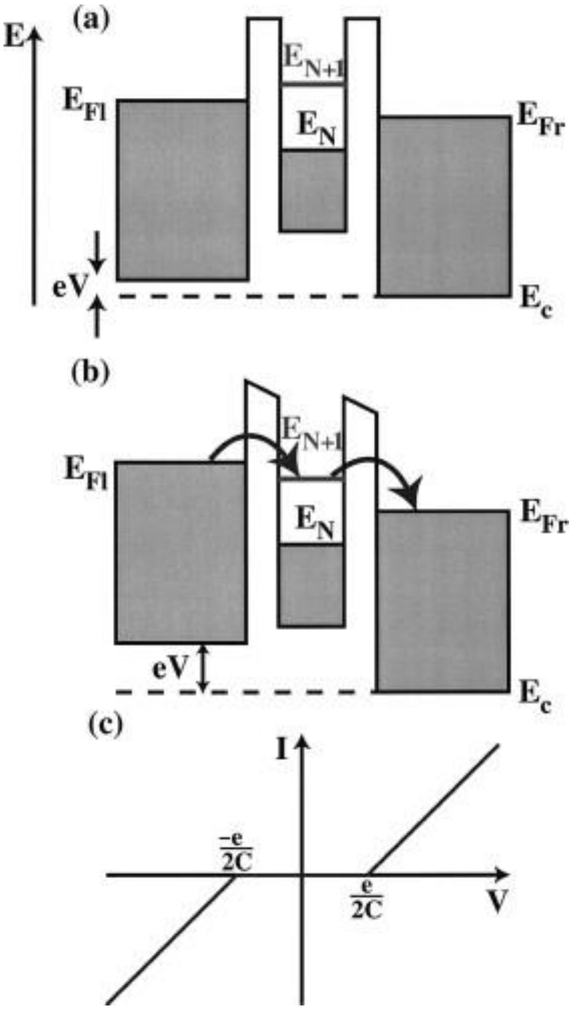


FIGURE 4. (a) The Coulomb blockade regime. An energy gap forms between the chemical potential E of the Nth and the $N + 1$th electrons. (b) Single electron tunneling. (c) The current voltage characteristics.

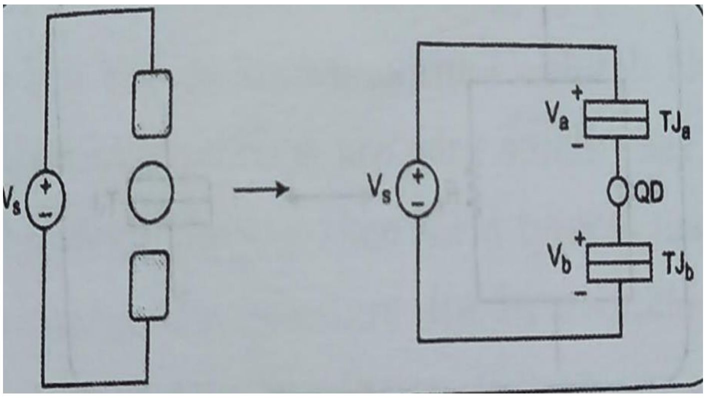
Condition for Coulomb Blockade

1.The bias voltage must be lower than the elementary charge divided by the self-capacitance of the island: V_{bias} ;< e/C

2.The thermal energy in the source contact plus the thermal energy in the island, i.e. KT must be below the charging energy KT< $e^2/2C$: or else the electron will be able to pass the QD via thermal excitation; and

3.The tunneling resistance, R_t should be greater than h/e^2 which is derived from Heisenberg's uncertainty principle.

Coulomb Blockade in a quantum dot circuit



Suppose a metallic dot is coupled to an external load as shown in fig. Tunnel junction represent the insulating regions isolating the quantum dot from the leads connected to external source V_s. Each of the tunnel junctions has been modelled by leaky capacitance $C_a$and resistance R_{ta} ; C_b, R_{tb} respectively. Associated Charge $Q_a = C_aV_a$and $Q_b = C_b V_b$, net charge of island is $Q= Q_a- Q_b$

The energy stored in the capacitors is

$$E_{sc}= \frac{Q_a^2}{2C_a} + \frac{Q_b^2}{2C_b} = \frac{1}{2} (C_aV_a^2 + C_bV_b^2) \qquad (1)$$

$$V_s = V_a + V_b= V_a + (Q+V_aC_a)/C_b \qquad (2)$$

As electron tunnelling allows discrete no electrons to accumulate on the island. Hence Q = nq , putting this in (2) we get

$$V_a = (V_sC_b - nq)/C_s\,, \quad C_s = C_a + C_b \qquad (3)$$

$$V_b = (V_sC_a + nq)/C_s \qquad (4)$$

As a result eq (1) become

$$E_{sc} = \frac{C_aC_b\,V_s^2 + (nq)^2}{2C_s} \qquad (5)$$

Suppose one electron tunnels onto the isl$(V_sC_b - nq)/C_s$ and through junction b. The voltage drops across the junction a and b will be respectively

$$V_a^f = [V_sC_b - (n+1)q]/C_s = V_a^i - \frac{q}{C_s}$$

$$V_b^f = [V_sC_a + (n+1)q]/C_s = V_b^i + \frac{q}{C_s} \qquad (6)$$

Such that $V_s = V_a + V_b$ is maintained. Consequently the resulting charge stored by the junction a is given by

$$Q_a^f = C_aV_a^f = C_a\left(V_a^i - \frac{q}{C_s}\right) = Q_a^i - C_aq/C_s \qquad (7)$$

So change in charge $\Delta Q = Q_a^i - Q_a^f = \frac{qC_a}{C_s}$, must come from power supply & associated with work W = $V_s\,\Delta Q$. Once the electron tunnels onto island through junction b, the change in total energy

$$\Delta E_t = \Delta E_{sc} - W$$

$$= \frac{C_aC_b\,V_s^2 - (nq)^2}{2C_s} - \frac{C_aC_b\,V_s^2 - \{(n+1)q\}^2}{2C_s} - V_s\,\frac{qC_a}{C_s}$$

$$= -\frac{q}{C_s}\left[\,q\,(n + ½) + V_s\,C_a\right] \qquad (8)$$

Now the tunnelling event is energetically favorable only when the above energy change is negative & consequently we get the condition as

$$V_s > -\,-\frac{q}{C_a}\left(n + \frac{1}{2}\right) \qquad (9)$$

Similarly when an electron tunnels off the island through junction a, the voltage across the junction will change as

$$V_a^f = [V_sC_b - (n-1)q]/C_s = V_a^i + \frac{q}{C_s} \qquad (10)$$

$$V_b^f = [V_sC_a + (n+1)q]/C_s = V_b^i - \frac{q}{C_s} \qquad (11)$$

Thus, the resulting charge stored in junction b is

$$Q_b^f = C_bV_b^f = C_b\left(V_b^i - \frac{q}{C_s}\right) = Q_b^i - C_bq/C_s \qquad (12)$$

Thus change in total energy

$$\Delta E_t = \Delta E_{sc} - W$$

$$= \frac{C_aC_b V_s^2 + (nq)^2}{2C_s} - \frac{C_aC_b V_s^2 + \{(n-1)q\}^2}{2C_s} - V_s \frac{qC_b}{C_s}$$

$$= -\frac{q}{C_s}[\, q\,(n - ½) - V_s C_b] \qquad (13)$$

Again the requirement of favorable energy, we have the condition

$$V_s > \frac{q}{C_b}\left(n - \frac{1}{2}\right) \qquad (14)$$

As a special $C_a = C_b = C$ and n=0 , both the eq (9) and (14) reduce to a simple condition

$$V_s > -\frac{q}{2C}$$

Now for opposite situation, where electron tunnel onto the island through junction a and off the island through junction b. For $C_a = C_b = C$

And n=0, one can get the condition as

$$V_s < \frac{q}{2C}$$

Therefore we have simple condition in this special condition

$$-\frac{q}{2C} < V_s < \frac{q}{2C} \text{ or } |V_s| > q/2C$$

Which is nothing but same Coulomb Blokade as encountered for single capacitor and is associated with the charging energy

$E_c = q^2/2C$

Now assume that one electron has already tunneled onto the island through junction b. If a second electron onto the dot through the junction b, the change in energy is

$$\Delta E_t = \Delta E_{sc} - W$$

$$= \frac{C_a C_b V_s^2 + \{(n+1)q\}^2}{2C_s} - \frac{C_a C_b V_s^2 + \{(n+2)q\}^2}{2C_s} - V_s \frac{qC_a}{C_s}$$

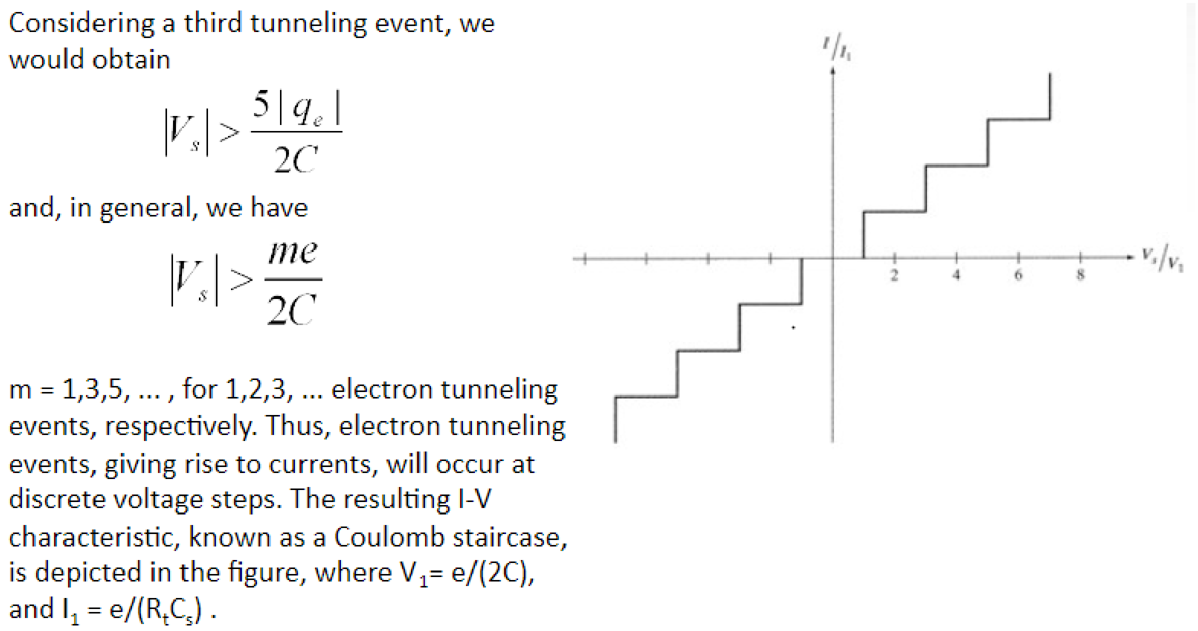
$$= -\frac{q}{C_s} [\, q\,(n + 3/2) + V_s C_a]$$

For favorable energetics ,we get condition

$$V_s > -\frac{q}{C_a} (n + 3/2)$$

For special case $C_a = C_b = C$ and n=0 and considering different tunnel direction we obtain

$$|V_s| > \frac{3q}{2C}$$



Considering a third tunneling event, we would obtain

$$|V_s| > \frac{5|q_e|}{2C}$$

and, in general, we have

$$|V_s| > \frac{me}{2C}$$

m = 1,3,5, ... , for 1,2,3, ... electron tunneling events, respectively. Thus, electron tunneling events, giving rise to currents, will occur at discrete voltage steps. The resulting I-V characteristic, known as a Coulomb staircase, is depicted in the figure, where $V_1 = e/(2C)$, and $I_1 = e/(R_t C_s)$.

blockade and staircase for a quantum dot

Variable-range hopping

Variable-range hopping is a model used to describe carrier transport in a disordered semiconductor or in amorphous solid by hopping in an extended temperature range. It has a characteristic temperature dependence of

$$\sigma = \sigma_0 e^{-(T_0/T)^{\beta}}$$

where σ is the conductivity and is a parameter dependent on the model under consideration.

Mott variable-range hopping

The **Mott variable-range hopping** describes low-temperature conduction in strongly disordered systems with localized charge-carrier states and has a characteristic temperature dependence of

$$\sigma = \sigma_0 e^{-(T_0/T)^{1/4}}$$

for three-dimensional conductance (with β = 1/4), and is generalized to *d*-dimensions

$$. \quad \sigma = \sigma_0 e^{-(T_0/T)^{1/(d+1)}} .$$

Hopping conduction at low temperatures is of great interest because of the savings the semiconductor industry could achieve if they were able to replace single-crystal devices with glass layers.[3]

In this section we continue our discussions of conduction in a 'Fermi glass', that is a degenerate electron gas in a highly disordered medium, and consider what happens to the d.c. conductivity when the Fermi energy lies in the range of energies where states are localized. Then two mechanisms for conduction are possible.

(i) Excitation of electrons to E_C; the contribution to the conductivity is

$$\sigma = \sigma_{\min} \exp\left(-\frac{\varepsilon}{kT}\right) \tag{2.57}$$

where $\sigma_{\min}$ is as before the value of σ_E at $E = E_C$, and

$$\varepsilon = E_C - E_F.$$

This form of conduction is normally predominant at high temperatures or when ε is small. In an alloy or other system in which the composition x changes in such a way that E_F passes through E_C, ε should tend linearly to zero, as already illustrated in Fig. 2.6(b).

(ii) Thermally activated hopping conduction by electrons in states near the Fermi energy. This is illustrated in Fig. 2.9; the rate-determining

(ii) Thermally activated hopping conduction by electrons in states near the Fermi energy. This is illustrated in Fig. 2.9; the rate-determining

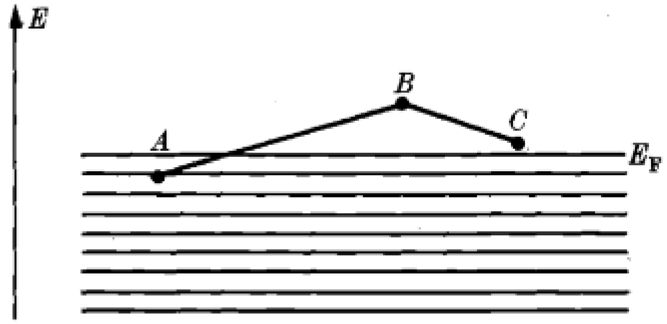


Fig. 2.9. The mechanism of hopping conduction. Two hops are shown, from A (an occupied state) to B and from B to C.

process is the hopping of an electron from a state (A) below the Fermi energy to one above (B). The probability p per unit time that this occurs is the product of the following factors.

(a) The Boltzmann factor $\exp(-W/kT)$, where W is the difference between the energies of the two states.

(b) A factor ν_{ph} depending on the phonon spectrum, discussed in Chapter 3.

(c) A factor depending on the overlap of the wavefunctions.

If localization is very strong (Fig. 2.2(d)), an electron will normally jump to the state nearest in space because the term $\exp(-2\alpha R)$ falls off rapidly with distance. This we call 'nearest-neighbour' or 'Miller–Abrahams' hopping, after the work of these authors (Miller and Abrahams 1960; cf. § 4.3) on impurity conduction. The conductivity is obtained as follows. The number of electrons jumping a distance R in the direction of the field will be made up of the following two factors.

(i) The number of electrons per unit volume within a range kT of the Fermi energy, namely $2N(E_F)kT$.

(ii) The difference of the hopping probabilities in the two directions, which are

$$\nu_{\mathrm{ph}} \exp\left\{-2\alpha R - \frac{W \pm eRF}{kT}\right\}$$

where F is the field. The current j is obtained by multiplying by e and R and is thus

$$j = 2eRkTN(E_F)\nu_{\mathrm{ph}} \exp(-2\alpha R - W/kT)\sinh(eRF/kT). \qquad (2.58)$$

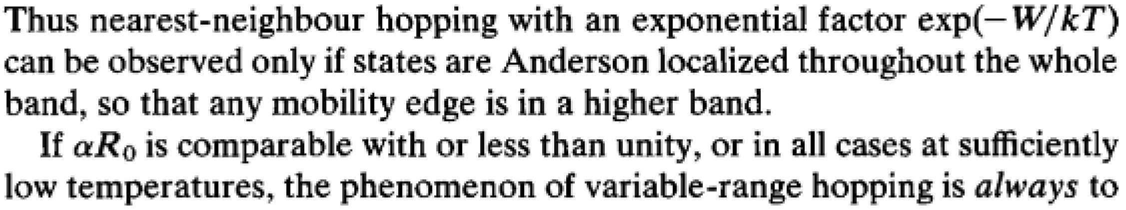
For weak fields, $eRF \ll kT$, the conductivity is

$$\sigma = j/F = 2e^2R^2\nu_{\mathrm{ph}}N(E_F)\exp(-2\alpha R - W/kT). \qquad (2.59)$$

Nearest-neighbour hopping is only expected if $\alpha R_0 \gg 1$, where R_0 is the average distance to a nearest neighbour. The hopping energy W is of the order of the bandwidth; we write

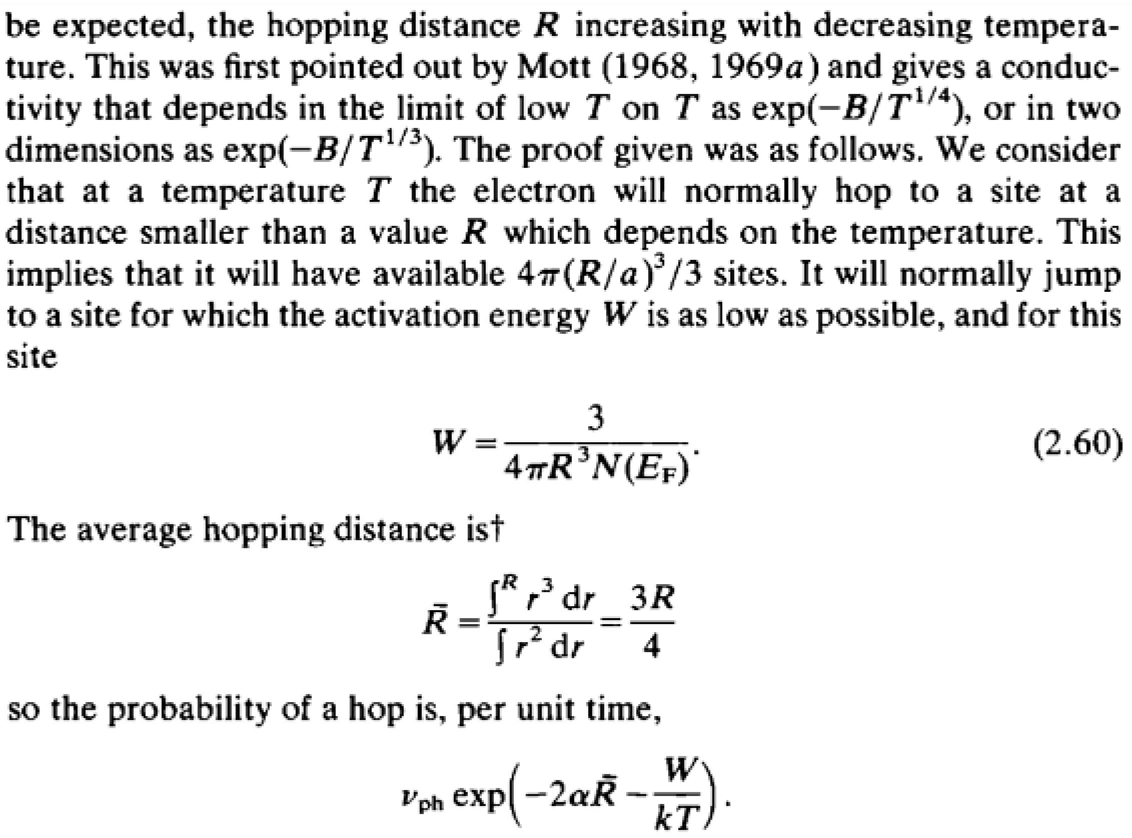
$$W \sim 1/R_0^3N(E_F).$$

[No of states within radius R is $N(E_F)$. So $(4\pi/3)\, R^3\, N(E_F)\, W(R) = 1$]



Thus nearest-neighbour hopping with an exponential factor $\exp(-W/kT)$ can be observed only if states are Anderson localized throughout the whole band, so that any mobility edge is in a higher band.

If αR_0 is comparable with or less than unity, or in all cases at sufficiently low temperatures, the phenomenon of variable-range hopping is *always* to



be expected, the hopping distance R increasing with decreasing temperature. This was first pointed out by Mott (1968, 1969*a*) and gives a conductivity that depends in the limit of low T on T as $\exp(-B/T^{1/4})$, or in two dimensions as $\exp(-B/T^{1/3})$. The proof given was as follows. We consider that at a temperature T the electron will normally hop to a site at a distance smaller than a value R which depends on the temperature. This implies that it will have available $4\pi(R/a)^3/3$ sites. It will normally jump to a site for which the activation energy W is as low as possible, and for this site

$$W = \frac{3}{4\pi R^3 N(E_F)}. \tag{2.60}$$

The average hopping distance is†

$$\bar{R} = \frac{\int^R r^3\,dr}{\int r^2\,dr} = \frac{3R}{4}$$

so the probability of a hop is, per unit time,

$$\nu_{ph} \exp\left(-2\alpha\bar{R} - \frac{W}{kT}\right).$$

Assuming γ_{ph} varies little with R and T, We have to take maximum value of this quantity, which occours when

P = γ_{ph} exp (-$2\alpha R_{av} - W/KT)$

$\frac{\partial P}{\partial R} = -\gamma_{ph}(2\alpha(\frac{3}{4}) - \frac{3\ x3}{4\pi R^4\ N(E_F)KT})$ exp(-$2\alpha R_{av} - 3/4\pi R^3 N(E_f\ )KT)$

(as R_{av} = $\frac{3}{4}$ R ; W = $\frac{3}{4\pi R^3 N(Ef)}$)

For maximum probability of hopping ($\partial P/\partial R$) = 0

$$\frac{3\alpha}{2} = \frac{9}{4\pi R^4 N(E_{f)}\ KT}$$

Giving optimum value of R

$$R = \frac{3^{1/4}}{(2\pi\alpha N(E_F)KT)^{1/4}} \qquad [A]$$

With the hopping distance 3/4 of this. The hopping probability therefore become

Hopping probability become

$P = \gamma_{ph} \exp(-2\alpha R_{av} - W/KT)$

[Putting value of R_{av}= 3/4 R = 3/4 $\left[\frac{3^{\frac{1}{4}}}{(2\pi\alpha N(E_F)KT)^{\frac{1}{4}}}\right]$ (from eq A)

& $W = \frac{3}{4\pi R^3 N(Ef)}$ we get]

$$P = \gamma_{ph} \exp\left[-2\alpha \ x \ \frac{3}{4}\left[\frac{3^{\frac{1}{4}}}{\{2\pi\alpha N(E_F)KT\}^{\frac{1}{4}}}\right] - \frac{3}{4\pi\left\{\frac{3}{2\pi\alpha N(E_F)KT}\right\}^{\frac{3}{4}} N(E_F)KT}\right]$$

$$= \gamma_{ph} \exp\left[-\frac{3\,\alpha^{3/4}3^{1/4}}{2\{2\pi N(E_F)KT\}^{1/4}} - \frac{\alpha^{3/4}3^{1/4}}{2\{2\pi N(E_F)KT\}^{1/4}}\right]$$

$$= \gamma_{ph} \exp\left[-\frac{\alpha^{3/4}3^{1/4}}{\{2\pi N(E_F)KT\}^{1/4}} \ (3/2 + 1/2)\right]$$

$$= \gamma_{ph} \exp -2\left[\frac{\alpha^{3/4}3^{1/4}}{\{2\pi N(E_F)KT\}^{1/4}}\right]$$

$$= \gamma_{ph} \exp -\left[2\left(\frac{3}{2\pi}\right)^{1/4}\left\{\frac{\alpha^3}{KN(E_F)}\right\}^{\frac{1}{4}}\left(\frac{1}{T}\right)^{\frac{1}{4}}\right]$$

$$= \gamma_{ph} \exp[-B/T^{1/4}]$$

With $B = B_o\left\{\frac{\alpha^3}{KN(E_F)}\right\}^{\frac{1}{4}}$ & $B_o = 2\left(\frac{3}{2\pi}\right)^{1/4}$

$$P = \gamma_{ph} \exp[-B/T^{1/4}] \qquad (B)$$

conductivity is obtained by multiplying (B) by $e^2N(E_F)(Rav)^2$

therefore conductivity is given by

$$\sigma = \quad e^2N(E_F)(R_{av})^2 \; \gamma_{ph} \; \exp\,[-B/T^{1/4}] \;\;)$$

$$= \quad e^2N(E_F) \;\; (3/4) \;\; \left[\frac{3^{\frac{1}{4}}}{(2\pi\alpha N(E_F)KT)^{\frac{1}{4}}}\right]^{\wedge}2 \; \exp\;[-B/T^{1/4}]$$

$$\boldsymbol{\sigma \sim \sigma_o \;\; T^{-1/2} \exp\,[-B/T^{1/4}]}$$

Validity of Mott’s Law

1. Temperature must be low enough. At high temperature, most probable hopping distance is comparable to nearest neighbor hopping distance and W(R) is independent of temp.For validity of Mott’s law W>KT
2. Density of state at the fermi level $N(E_F)$ should be non-zero and constant between two localized states E_F and $E_F \pm W$
   $\alpha R \gg 1, \quad KT \ll E_F$
3. Tunneling factor is explicitly dictated by exponential decay of wave function.
4. Multi -Phonon process have not taken into consideration.
5. Correlation effects in tunnelling states and electron-phonon interaction habe been ignored.

What is Crystal Defect?

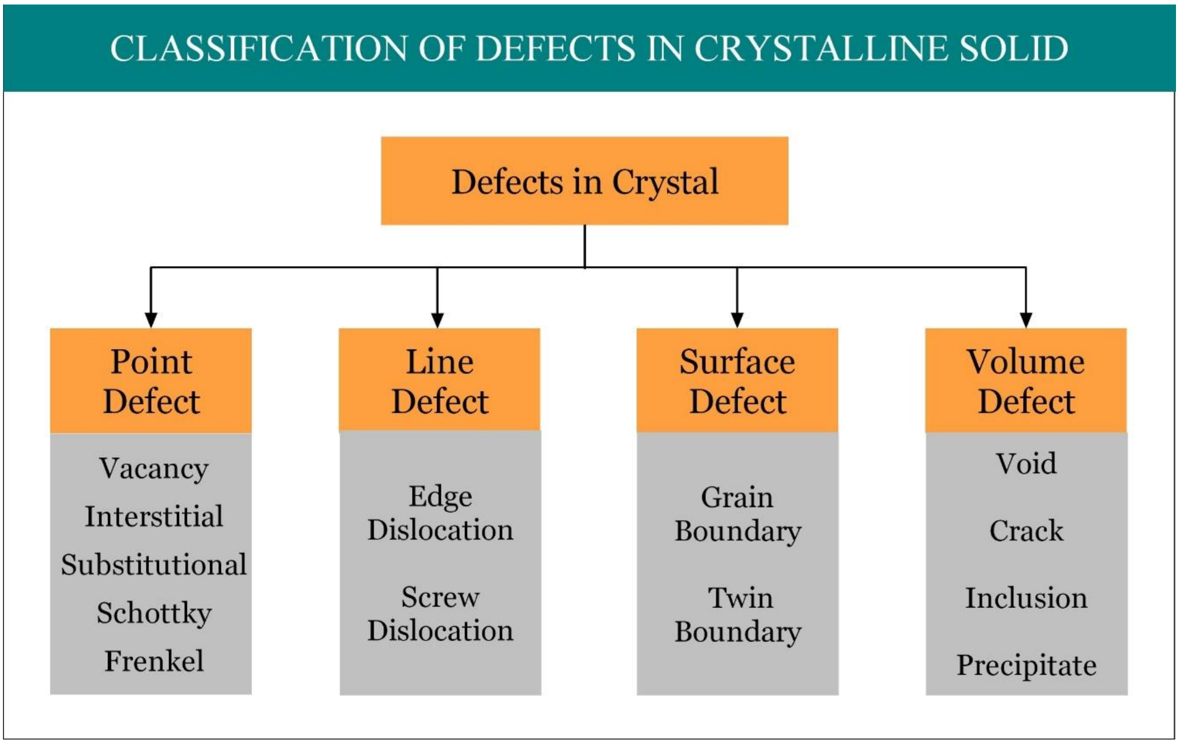
An ideal crystal can be described in terms of a three-dimensionally periodic arrangement of points called lattice and an atom or group of atoms associated with it.

Crystal lattice = Lattice + Basis

where the lattice is a 3D periodic arrangement of points and Basis is an atom or group of atoms. Any deviations from the perfect arrangement of atoms in a crystal are called crystal defects. A crystal defect is a lattice irregularity having one or more of its dimensions on the order of an atomic diameter. Imperfections have a significant impact on the characteristics of some materials. Because of this, it's critical to understand the different kinds of crystal defects that occur and how they affect how materials behave. For example, alloying significantly affects the mechanical characteristics of pure metals (i.e., when impurity atoms are added)—

e.g., sterling silver (92.5% silver-7.5% copper) is much harder and stronger than pure silver.

- Substitutional Defects(Impurity Defect)



Point Defects

Point defects (type of crystal defects) have atomic dimensions and occur only at or around a single lattice point. They are not stretched in any dimension in space. Different types of point defects in crystals are shown below -

- Stoichiometric Defect (Vacancy Defects, Interstitial Defects, Frenkel Defects, Schottky Defects)
- Substitutional Defects(Impurity Defect)
- Non Stoichiometric Defect

Stoichiometric Defect:

In this kind of point defect, the ratio of positive and negative ions (Stoichiometric) and electrical neutrality of a solid is not disturbed. Sometimes it is also known as intrinsic or thermodynamic defects.

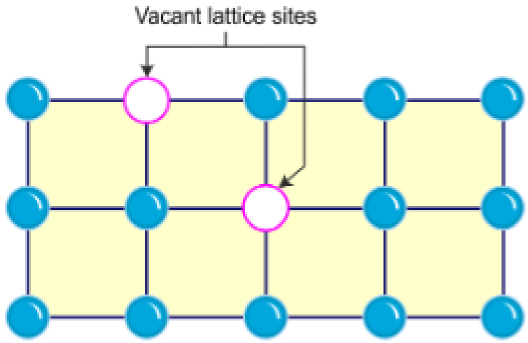
Fundamentally, they are of two types:

- ***Vacancy defect:*** When an atom is not present at their lattice sites, then that lattice site is vacant and it creates a vacancy defect. Due to this, the density of a substance decreases.
- ***Interstitial defect:*** It is a defect in which an atom or molecule occupies the intermolecular spaces in crystals. In this defect, the density of the substance increases.

A non-ionic compound mainly shows vacancy and interstitial defects. An ionic compound shows the same in Frenkel and Schottky defect.

Vacancy Defects

A missing atom at a lattice position causes a vacancy defect. The surrounding crystal structure's stability ensures that the atoms do not just collapse around the void. The vacancy kind of defect can be caused by poor packing during the crystallisation process, or by enhanced thermal vibrations of the atoms caused by high temperatures.



Vacancy defect

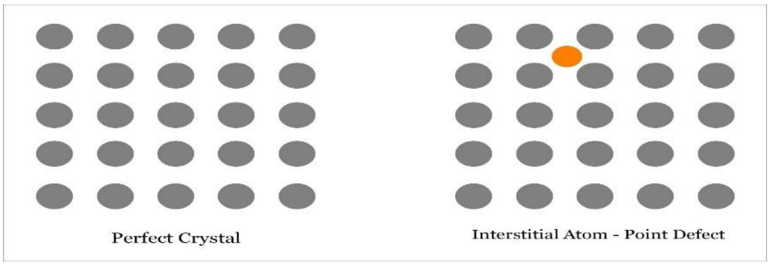
Interstitial – A Point Defect

An interstitial defect occurs when an atom takes the interstitial position of the lattice structure. This interstitial atom may be of the same crystal or of a foreign material. Accordingly, interstitial defect can be of two types:

- **Self-Interstitial Defect**—occurs when atom of the same crystalline solid occupies the interstitial position leaving its original lattice site.

- **Interstitial Defect**—occurs when a foreign atom occupies the interstitial position.

Although extra atom occupies the empty interstitial space, the size of the atom is usually larger than that of the empty space. Thus the surrounding atoms are compressed and distorted. Presence of substantial number of interstitial atoms can change the mechanical and thermal properties of the solid. However, this is sometime beneficial, and thus interstitial defects can be applied in a controlled way to enhance various properties of the solid. For example, in steel production, carbon is added with iron.



Frenkel Defect:

Frenkel defect is a type of point defect in crystalline solids, named after its discoverer Yakov Frenkel. The defect forms when an atom or smaller ion (usually cation) leaves its place in the lattice, creating a vacancy and becomes an interstitial by lodging in a nearby location.

In elemental systems, they are primarily generated during particle irradiation, as their formation enthalpy is typically much higher than for other point defects, such as vacancies, and thus their equilibrium concentration according to the Boltzmann distribution is below the detection limit.

In ionic crystals, which usually possess low coordination number or a considerable disparity in the sizes of the ions, this defect can be generated also spontaneously, where the smaller ion (usually the cation) is dislocated. Similar to a Schottky defect the Frenkel defect is a stoichiometric defect (does not change the over all stoichiometry of the compound). In ionic compounds, the vacancy and interstitial defect involved are oppositely charged and one might expect them to be located close to each other due to electrostatic attraction. However, this is not likely the case in real material due to smaller entropy of such a coupled defect, or because the two defects might collapse into each other. Also, because such coupled complex defects are stoichiometric, their concentration will be independent of chemical conditions

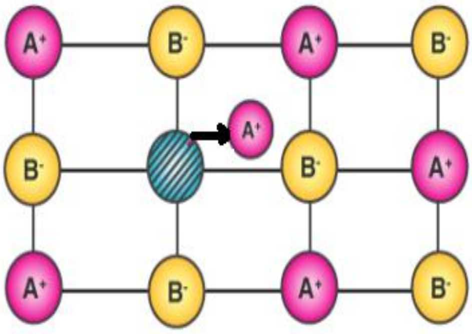
Even though Frenkel defects involve only the migration of the ions within the crystal, the total volume and thus the **density i**s not necessarily changed

The number of Frenkel Defects can be calculated using the equation:

$N=\sqrt{NN^*}\ \exp(-E_i/KT)$

where N is the number of normally occupied positions, N^* is the number of available positions for the moving ion, the E_i is the energy required to displace an atom/ion from its proper position to an interstitial position for formation of one Frenkel defect. Frenkel defects are intrinsic defects because the existence causes the Gibbs energy of a crystal to decrease, which means it is favorable to occur.

- It is also known as dislocation defect.
- The density of a substance remains unchanged.
- It happens when there is a huge difference in the size of anions and cations.
- ***Example:*** ZnS and AgCl



Frenkel Defect

© Byjus.com

Schottky Defect

A **Schottky defect** is an excitation of the site occupations in a crystal lattice leading to point defects named after Walter H. Schottky. In ionic crystals, this defect forms when oppositely charged ions leave their lattice sites and become incorporated for instance at the surface, creating oppositely charged vacancies. These vacancies are formed in stoichiometric units, to maintain an overall neutral charge in the ionic solid.

Typically, the formation volume of a vacancy is positive: the lattice contraction due to the strains around the defect does not make up for the expansion of the crystal due to the additional number of sites. Thus, the density of the solid crystal is less than the theoretical density of the material.

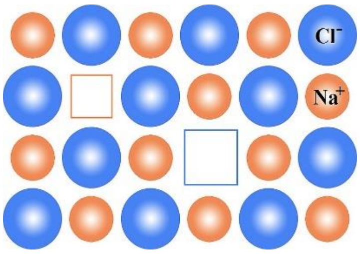
This type of defect is typically observed in highly ionic compounds, highly coordinated compounds, and where there is only a small difference in sizes of cations and anions of which the compound lattice is composed. Typical salts where Schottky disorder is observed are NaCl, KCl, KBr, CsCl and AgBr. For engineering applications, Schottky defects are important in oxides with Fluorite structure, such as CeO_2, cubic ZrO_2, UO_2, ThO_2 and PuO

It is possible to approximate the number of Schottky defects (n) in a MX ionic crystal compound by using the equation

$$n = N \exp\left(-\frac{E}{KT}\right)$$

E is energy required to take an atom from a lattice site inside the crystal to a lattice site on the surface.

- This kind of vacancy defects is found in Ionic Solids. But in ionic compounds, we need to balance the electrical neutrality of the compound so an equal number of anions and cations will be missing from the compound.
- It reduces the density of the substance.
- In this, the size of cations and anions are of almost the same.
-



Schottky Defect	Frenkel Defect
Schottky defect occurs in those ionic crystals where difference in size between cation and anion is small.	Frenkel defect usually occurs in those ionic crystals where size of anion is quite large as compared to that of the cation.
In Schottky defect, both cation and anion leave the solid crystal.	In Frenkel defect, only the smaller ion (cation) leaves its original lattice site; whereas, the anion remains in original lattice sites.
The atoms permanently leave the crystal.	Here, atoms leave the original lattice site and occupy interstitial position. So atoms reside within the solid crystal.
One Schottky defect leads to the formation of two vacancies.	One Frenkel defect creates one vacancy and one self-interstitial defect.
Two atoms reduce from the crystal for each Schottky defect.	The number of atoms present in the crystal before and after Frenkel defect remains same.
Due to vacancy formation, Schottky defect reduces density of the solid.	Density of the solid crystal before and after Frenkel defect remains same as no atom leaves the solid.
Common materials where Schottky defect can be found are: • Sodium Chloride (NaCl) • Potassium Chloride (KCl) • Potassium Bromide (KBr) • Silver Bromide (AgBr)	Common materials where Frenkel defect can be found are: • Zinc Sulfide (ZnS) • Silver Chloride (AgCl) • Silver Bromide (AgBr)

• Cerium Dioxide (CeO_2) • Thorium Dioxide (ThO_2)	

1. **Impurity Defect:** Let's understand the impurity defect by an example. If molten NaCl is crystallized with $SrCl_2$ compound then the Sr^{2+} ions replace two Na^+ ions and occupy the place of one Na^+ In this way the lattice site of one Na^+ is vacant and it creates an impurity defect.
2. **Non-Stoichiometric Defect:** In this defect, the cations and anions ratio is disturbed either because of adding or removing of ions.

Types of Non-Stoichiometric Defect:

1. ***Metal deficiency defect:*** In this, the solids have less number of metals relative to the described Stoichiometric proportion. The less proportion of metal is compensated by same metals having higher valance. They are generally observed in transition element.

Example: FeO is generally found in composition of $FeO_{.95}O_{0.05}$.The non-stoichiometry reflects the case of oxidation of Fe^{++} to Fe^{+++} effectively replacing a small portion of Fe^{++} with two third their no of Fe^{+++}. $3Fe^{+2}= 2Fe^{+3}$, maintains the electrical neutrality in the system

2. ***Metal excess defect:*** There are two types of metal excess defect:

- *Metal excess defect due to anionic vacancies*: This occurs due to the absence of anions from its original lattice site in crystals. Therefore, instead of anions, electrons occupy their position. Example : NaCl, KCl, LiCl. The anionic site occupied by unpaired electron is called F -centre.When visible light falls over NaCl, the unpaired electrons present get excited because of absorption of energy and impart yellow color.Because of this unpaired electron, msgnetic response of F-center is paramagnetic & susceptibility obeys Curie's law.

  Because of similar defect LiCl imparts pink color & KCl imparts violet.

- *Metal excess defect due to the presence of extra cations at interstitial sites*: Here, on heating the compound, it releases extra cations. These cations occupy the interstitial sites in crystals and the same number of electrons goes to neighbouring interstitial sites.

  ZnO→ Zn^{++} + ½ O_2 $+2e^-$

On heating the formula of ZnO becomes $Zn_{1+x}O$.

The excess cations move to interstitial site and electron move to neighboring interstitial site Because of the ZnO imparts yellow color.

Line Defects

Line defects (type of crystal defects) typically span a large number of atoms. Dislocations are line defects that only appear in crystalline materials. A dislocation is a linear or one-dimensional imperfection in which certain atoms are misaligned.

Dislocations are particularly essential in materials science because they contribute to material mechanical strength. The two fundamental forms of dislocations are -

- Edge dislocation line
- Screw dislocation

.

Difference between Edge dislocation and Screw dislocation:

Edge dislocation:

- Edge dislocation occurs due to the introduction or elimination of an extra row of atoms.
- Shear, Tensile and Compressive stress fields may be present in the Edge dislocation.
- In this dislocation, Burger’s vectors always perpendicular to the dislocation line.
- Region of lattice disturbance extends along an edge inside a crystal.
- This dislocation occurs due to climb and glide motion.
- There are two types of Edge dislocation, positive and negative.

Screw dislocation:

- Screw dislocation provides easy crystal growth because of additional unit cells and atoms can be added to the screw.
- Shear, Tensile and Compressive stress fields are absent in the Screw dislocation.

- In this dislocation, Burger’s vectors always parallel to the dislocation line.
- Region of lattice disturbance is in two separate planes and cross each other at right angles.
- This dislocation occurs only due to glide motion.
- Screw dislocation does not show any type like edge dislocation.

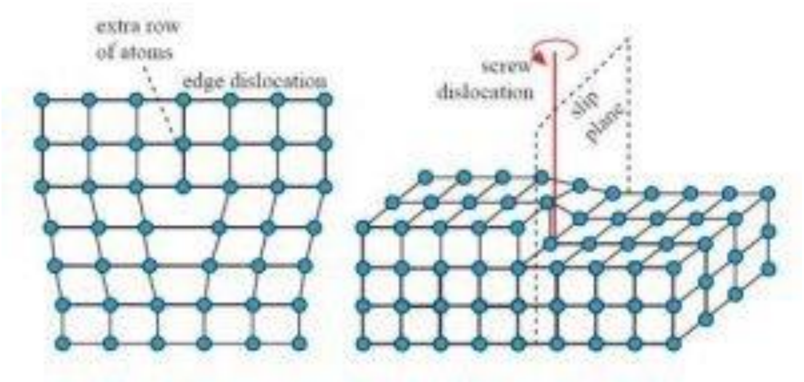


Diagram of Edge dislocation and Screw dislocation

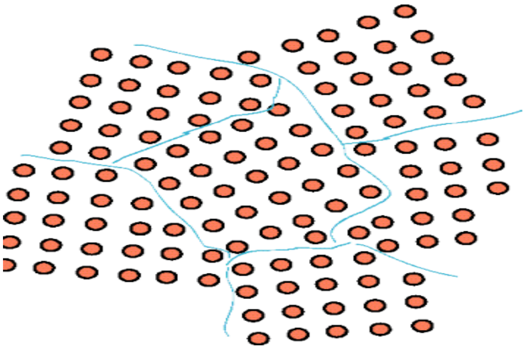
Planar or Interfacial Defects

Interfacial defects (type of crystal defects) are two-dimensional barriers that generally separate sections of materials with various crystal structures and/or crystallographic orientations. Different types of planar defects are -

- Grain Boundaries
- Twinning

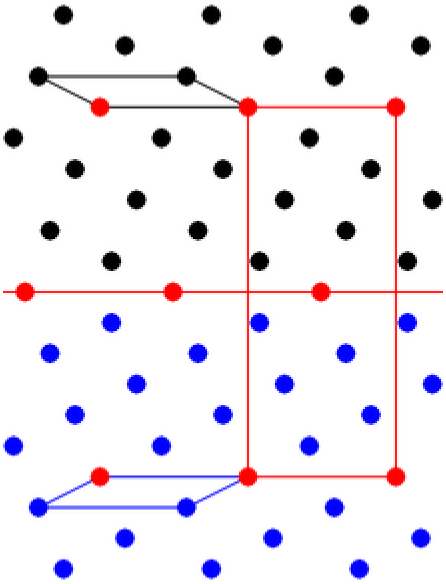
Grain Boundaries

Planar Defects: Grain Boundaries A Grain Boundary is a general planar defect that separates regions of different crystalline orientation (i.e. grains) within a polycrystalline solid. The atoms in the grain boundary will not be in perfect crystalline arrangement. Grain boundaries are usually the result of uneven growth when the solid is crystallising. Grain sizes vary from 1 µm to 1 mm.



Twinning

A twin boundary is a form of grain boundary that has mirror lattice symmetry; that is, atoms on one side of the boundary are in mirror image locations of atoms on the other side. It is correct to refer to the material region between these limitations as a twin.



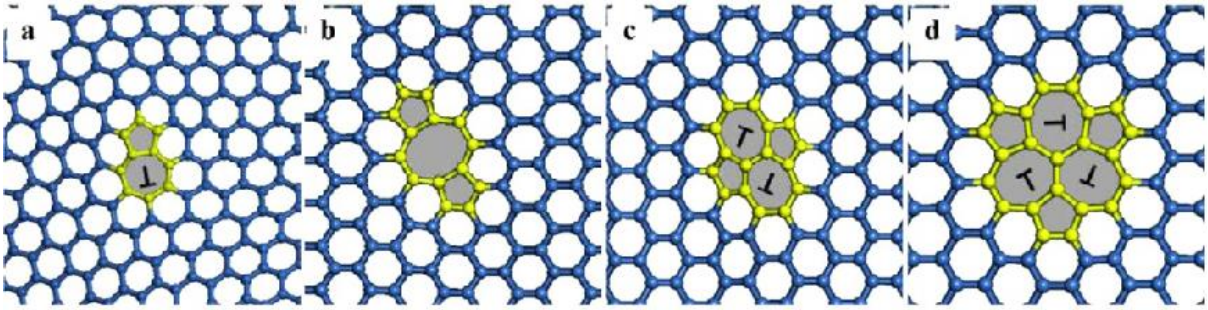
Twin lattice (2D). The long horizontal red line is the composition plane where the two crystal segments join. The upper crystal lattice is the reflection of the lower crystal lattice. The red points are the shared crystal lattice points.

Topological Defect

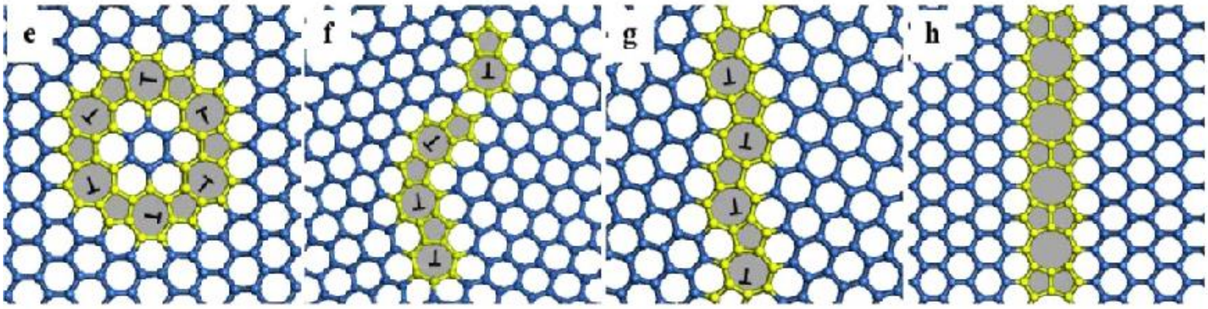
A topological defect in a material occours as a discontinuity in the order,ie a boundary region where the order does not seamlessly make a transition from one area to another.These topological structure form naturally or can be highly engineered in advanced functional material.These distortion generally arise from imposition of boundary conditions,from external fields,or from thermal fluctuations.These topological defects originate in system with broken continuous symmetry.

Topological defect play a key role in determining & controlling properties of raw material .They are responsible to a large degree for the mechanical properties of metals like steel.They are particularly important in two dimention, where they play a pivotal role in the transition from low temperature phase characterised by non vanishing rigidity to a high temperature disordered phase.The liquid crystal topological defect may be used as a marrix for new kinds of nano particle organizations teplate by defect geometry.

These defect change the electronic structure and suceptibilityto chemicalreaction which modifies the chemical reactivity of graphene.Density of states are also modified by these defect resulting changes in transport properties of the defected graphene system.It may not be ideal for some electronic applications but it can be used as a safe and biocompatible biomaterial in various biological systems.



5-7 dislocation　5-8-5 dislocation　TD55-77　TD555-777



Flower defect　Aperiodic GB　Periodic GB ($\theta = 32.2°$)　Periodic GB ($\theta = 0°$)

. Structures of topological defects (TD) in graphene: (a) 5 7 dislocation; (b) 5 8 5 dislocation; (c) 55 77 clustered topological defect; (d) 555 777 clustered topological defect; (e) flower defect, which is a closed loop of 5 7 pairs. (f) Aperiodic grain boundary (GB); (g) θ = 32.2 periodic grain boundary; (h) θ = 0 periodic translational grain boundary. Collapse

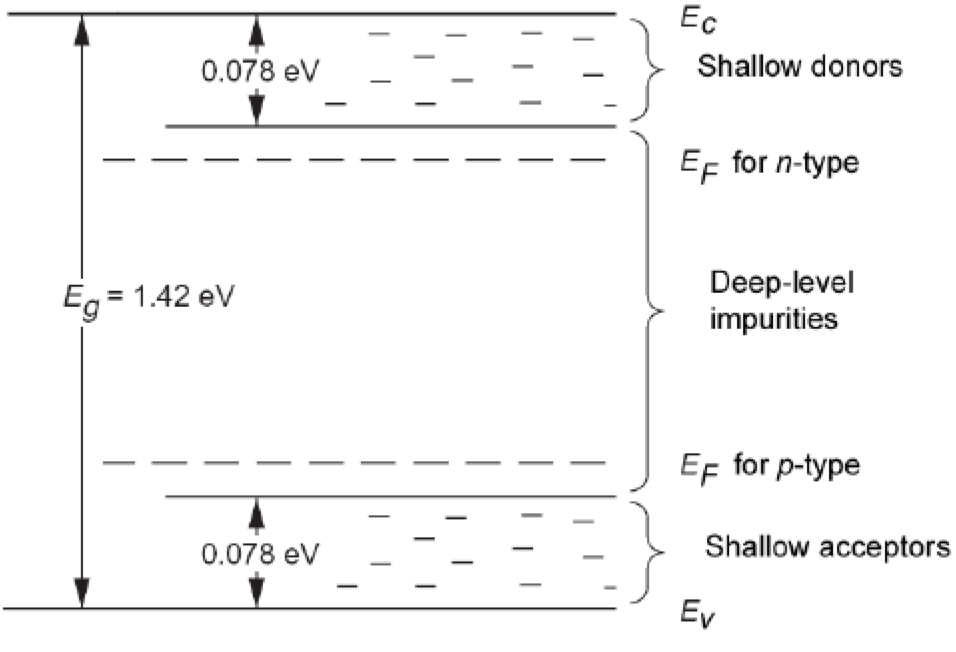
Deep level defects & impurities

In Deep level defect or Deep level trap energy required to remove an electron or hole from the trap to the valence or conduction band is much larger than characteristic thermal energy KT. Most often deep signifies the energy of the order of 150 meV from the conduction band or valence band edges.

The binding energies E of shallow impurities or shallow levels of donor and acceptor impurities are<100 meV. $E<<E_g$.

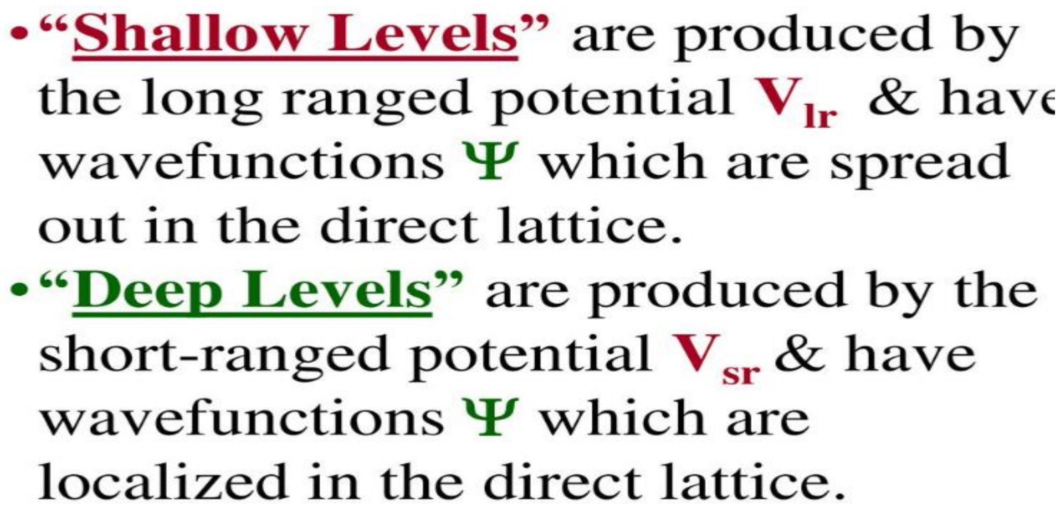
Characteristics features of deep level defects

1. Energy level E produced by deep centers may have energies in the band gap which can be close to either the conduction band edge or valence band edge.
2. For such defect lattice distortion effect are important
3. In Effective medium theory(EMT) , defect potential has form: V= $V_{lr}+V_{sr}$ where V_{lr} is long range Coulomb & V_{sr}short range potential due to impurity. EMT work well for Shallow level & fail for Deep level
4. Deep level can act as center for e^- - e^+ recombination.They can have strong effect on material optical and electronic properties.
5. Deep level can enhance the radiative recombination. They can be used in LED for specific color.
6. Deep level can shorten e^- and e^+ lifetimes. This is good for fast switch and bad for photo cell.



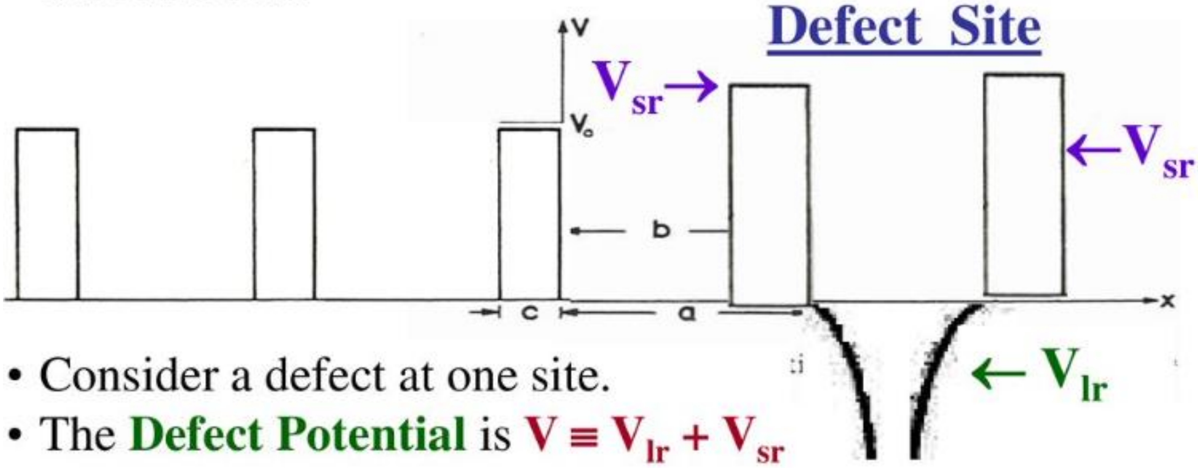
Comparison of Deep & Shallow level defect

Deep Level Defect	Shallow level defect
Concentration of defect producing deep level is 10^{12} to 10^{18} cm^{-3}	Concentration of defect is higher $\sim 10^{18}$ cm^{-3}
They have negligible effect on electrical conductivity. Their concentration are too small to affect the electron number density n.	Donor / Acceptor control conductivity. Current density $j=nev=\sigma E$
In EMT deep levels are produced by short range V_{sr}	In EMT, V_{lr} responsible for the shallow- donor and acceptor level.
Electron wave function are localized in r space.	Electron wave function are hydrogen -like spread out in r space.



- "**Shallow Levels**" are produced by the long ranged potential $\mathbf{V_{lr}}$ & have wavefunctions Ψ which are spread out in the direct lattice.
- "**Deep Levels**" are produced by the short-ranged potential $\mathbf{V_{sr}}$ & have wavefunctions Ψ which are localized in the direct lattice.

- Schematically, it helps to think about a one-dimensional model lattice, such as the **Krönig-Penney Model** from our bandstructure discussions:



- Consider a defect at one site.
- The **Defect Potential** is $\mathbf{V \equiv V_{lr} + V_{sr}}$