Introduction

Semiconductors

- Conductivity in between those of metals and insulators.
- Conductivity can be varied over orders of magnitude by changes in temperature, optical excitation, and impurity content (doping).
- Generally found in column IV and neighboring columns of the periodic table.
- Elemental semiconductors: Si, Ge.
- Compound semiconductors:

  Binary:

  - GaAs, AlAs, GaP, etc. (III-V).
  - ZnS, ZnTe, CdSe (II-VI).
  - SiC, SiGe (IV compounds).

  - Ternary: GaAsP.

  Quaternary: InGaAsP.

- Si widely used for rectifiers, transistors, and ICs.
- III-V compounds widely used in optoelectronic and high-speed applications.

Applications

- Integrated circuits (ICs) SSI, MSI, LSI, and VLSI.
- Fluorescent materials used in TV screens II-VI (ZnS).
- Light detectors InSb, CdSe, PbTe, HgCdTe.
- Infrared and nuclear radiation detectors Si and Ge.
- Gunn diode (microwave device) GaAs, InP.
- Semiconductor LEDs GaAs, GaP.
- Semiconductor LASERs GaAs, AlGaAs.

Energy Gap

- Distinguishing feature among metals, insulators, and semiconductors.
- Determines the absorption/emission spectra, the leakage current, and the intrinsic conductivity.
- Unique value for each semiconductor (e.g. 1.12 eV for Si, 1.42 eV for GaAs) function of temperature.

Impurities
• Can be added in precisely controlled amounts.
• Can change the electronic and optical properties.
• Used to vary conductivity over wide ranges.
• Can even change conduction process from conduction by negative charge carriers to positive charge carriers and vice versa.
• Controlled addition of impurities doping.

[toggle]

Crystal Lattices

• Semiconductor properties can be strongly affected by crystal structure.

Types of Solids

• Basically, there are three types of solids: crystalline, amorphous, and polycrystalline.

Crystaline Solids

• Atoms making up the solid arranged in a periodic fashion, repeated throughout.
• Have long-range order.
• Used for IC fabrication.

Amorphous Solids

• Have no periodic structure at all.
• Interatomic distance and bond angles are almost the same as in the crystalline material of the same substance, however, a long-range order is missing.
• Said to have short-range order.
• a-Si (alloy of amorphous Si with H2 and other similar amorphous alloys) has found important applications in photovoltaic technology and in large-area ICs used in flat displays, printers, copiers, scanners, and imagers.

Polycrystalline Solids

• Composed of many small regions of single-crystal material of irregular size, separated by grain boundaries.

Lattice

• 3-D periodic arrangement of atoms in a crystal.
• Defined by primitive basis vectors \( \vec{a}, \vec{b}, \text{and} \vec{c} \), which are three independent shortest vectors connecting lattice sites.
• The coordinates of all points belonging to the crystal lattice are given by vectors 
\[
\vec{R}_{k,l,m} = k\vec{a} + l\vec{b} + m\vec{c}
\]
, where \( k \), \( l \), and \( m \) are integers.
• Properties of the periodic crystal determine the allowed energies of electrons that participate in the conduction process. Thus, the lattice not only determines the mechanical properties of the crystal, but also its electrical properties.

❖ Unit Cell

• Representative of the entire crystal and regularly repeated throughout the crystal.
• The crystal can be analyzed as a whole by investigating a representative volume.
• Can find:
  i) the distances between nearest atoms and next nearest atoms,
  ii) the fraction of the unit cell volume filled by atoms, and
  iii) the density of the solid (related to the atomic arrangement).

❖ Primitive Cell

• Smallest unit cell that can be repeated to form the lattice.

❖ Cubic Lattices

• Simplest 3-D lattice, where the unit cell is a cube.
• Three types:
  i) simple cubic (sc) (e.g., Ga),
  ii) body-centered cubic (bcc) (e.g., Na, W), and
  iii) face-centered cubic (fcc) (e.g., Al, Au).
• Lattice constant: the length of each side of the cube.

❖ The Diamond Lattice

• Two interpenetrating fcc sublattices spaced 1/4th along the body diagonal.
• When the constituent atoms of the two sublattices are different, then the structure is Zincblende (e.g. GaAs).
Diamond and Zincblende are the two most common crystal structures for cubic semiconductors.

Each atom in diamond and zincblende lattice is surrounded by four nearest neighbors.

Tetrahedral configuration.

By varying the atomic compositions of these two sublattices, one can grow ternary (e.g. $Al_xGa_{1-x}As$) and quaternary (e.g., $In_xGa_{1-x}As_yP_{1-y}$) compounds.

**EXAMPLE 1.1:** Find the fraction of the unit cell volume filled with hard spheres for a diamond lattice.

**SOLUTION:** Diamond lattice unit cell consists of an fcc lattice along with 4 inside atoms. The corner atoms in an fcc lattice is shared by seven nearest neighbors, thus each contributing $1/8$th of a sphere, and the face atoms are shared by one nearest neighbor, thus each contribution $1/2$ of a sphere.

Thus, $\text{Atoms/cell} = 1 \text{ (corners)} + 3 \text{ (faces)} + 4 \text{ (inside)} = 8$ Nearest neighbor distance = $(a\sqrt{3})/4$

Radius of each sphere = $(a\sqrt{3})/8$ Volume of each sphere = $(4/3)\pi[(a\sqrt{3})/8]^3$

\[
\text{Maximum fraction of cell filled} = \frac{\text{no. of spheres } \times \text{ vol. of each sphere}}{\text{total vol. of cell}} = \frac{4/3 \pi [(a\sqrt{3})/8]^3}{a^3} = 0.34
\]

Therefore, the unit cell volume filled with hard spheres for a diamond lattice is only 34%, thus, the diamond lattice is relatively loosely packed.
Planes and Directions

Miller Indices

- Found by taking the reciprocal of the intercepts of the plane with the coordinate axes, and converting these to integers.
- Define a set of parallel planes.
  - (hkl) a particular plane.
  - {hkl} equivalent (hkl) planes.
  - [hkl] a particular direction, perpendicular to the (hkl) plane for cubic crystals.
  - <hkl> equivalent [hkl] directions.
- Negative intercepts/directions are denoted by placing bars above the integers, e.g. \( \bar{h} \bar{k} \bar{l} \).
- Three main indices (100), (110), and (111).
- Atom densities different for different planes different electronic properties.

EXAMPLE 1.2: Show a (643) plane and (643) direction in a cubic lattice.

SOLUTION: (643) plane: reciprocals (1/6,1/4,1/3), converting them to smallest set of integers gives (2,3,4); thus, the plane has (x,y,z) intercepts as (2,3,4).

Practice Problems

1.1 Determine the nearest neighbor distance, and the number of the nearest and the next neighbors of sc, bcc, fcc, and diamond lattices with lattice constant a.

1.2 Determine the maximum fractions of the unit cell volume that can be filled by hard spheres in the sc, bcc, and fcc lattices.

1.3 Clearly state the equivalent \{100\}, \{110\}, and \{111\} planes and determine their total numbers in a cubic crystal.
1.4 Show the (211) plane and the [211] direction in a cubic crystal lattice.

1.5 Calculate the densities of Ge (diamond structure, $a = 5.66 \text{ Å}$) and InP (Zincblende structure, $a = 5.87 \text{ Å}$). The atomic weights of Ge, In, and P are 72.6, 114.8, and 31 respectively.

1.6 Determine the atom density (in number of atoms/cm$^2$) in (100), (110), and (111) planes of Si ($a = 5.43 \text{ Å}$).

Energy Bands and Charge Carriers in Semiconductors

**Bonding Forces and Energy Bands in Solids**

- Electrons are restricted to sets of discrete energy levels within atoms, with large gaps among them where no energy state is available for the electron to occupy.
- Electrons in solids also are restricted to certain energies and are not allowed at other energies.
- Difference in the solid, the electron has a range (or band) of available energies.
- The discrete energy levels of the isolated atom spread into bands of energies in the solid because
  i) in the solid, the wave functions of electrons in neighboring atoms overlap, thus, it affects the potential energy term and the boundary conditions in the Schrödinger equation, and different energies are obtained in the solution, and
  ii) an electron is not necessarily localized at a particular atom.
- The influence of neighboring atoms on the energy levels of a particular atom can be treated as a small perturbation, giving rise to shifting and splitting of energy states into energy bands.

**Bonding Forces in Solids**

**Ionic Bonding**

- Example: NaCl.
- Na ($Z = 11$) gives up its outermost shell electron to Cl ($Z=17$) atom, thus the crystal is made up of ions with the electronic structures of the inert atoms Ne and Ar.
- Note: the ions have net electric charges after the electron exchange $\rightarrow$ Na$^+$ ion has a net positive charge, having lost an electron, and Cl$^-$ ion has a net negative charge, having acquired an electron.
- Thus, an electrostatic attractive force is established, and the balance is reached when this equals the net repulsive force.
- Note: all the electrons are tightly bound to the atom.
- Since there are no loosely bound electrons to participate in current flow $\rightarrow$ NaCl is a good insulator.
Metallic Bonding

- In metals, the outer shell is filled by no more than three electrons (loosely bound and given up easily) \( \Rightarrow \) great chemical activity and high electrical conductivity.
- Outer electron(s) contributed to the crystal as a whole \( \Rightarrow \) solid made up of ions with closed shells immersed in a sea of free electrons, which are free to move about the crystal under the influence of an electric field.
- Coulomb attraction force between the ions and the electrons hold the lattice together.

Covalent Bonding

- Exhibited by the diamond lattice semiconductors.
- Each atom surrounded by four nearest neighbors, each having four electrons in the outermost orbit.
- Each atom shares its valence electrons with its four nearest neighbors.
- Bonding forces arise from a quantum mechanical interaction between the shared electrons.
- Both electrons belong to each bond, are indistinguishable, and have opposite spins.
- No free electrons available at 0 K, however, by thermal or optical excitation, electrons can be excited out of a covalent bond and can participate in current conduction \( \Rightarrow \) important feature of semiconductors.

Mixed Bonding

- Shown by III-V compounds bonding partly ionic and partly covalent.
- Ionic character of bonding becomes more prominent as the constituent atoms move further away in the periodic table, e.g., II-VI compounds.

Energy Bands

- As isolated atoms are brought together to form a solid, the electron wave functions begin to overlap.
- Various interactions occur, and, at the proper interatomic spacing for the crystal, the forces of attraction and repulsion find a balance.
- Due to Pauli exclusion principle, the discrete energy levels of individual atoms split into bands belonging to the pair instead of to individual atoms.
- In a solid, due to large number of atoms, the split energy levels for essentially \textit{continuous bands of energy}. 


Fig.2.1 Splitting of individual energy levels to energy bands as atoms are brought closer together.

- Imaginary formation of a diamond crystal from isolated carbon atoms ($1s^22s^22p^2$).
- Each atom has two 1s states, two 2s states, six 2p states, and higher states.
- For $N$ atoms, the numbers of states are $2N$, $2N$, and $6N$ of type 1s, 2s, and 2p respectively.
- With a reduction in the interatomic spacing, these energy levels split into bands, and the 2s and 2p bands merge into a single band having $8N$ available states.
- As the interatomic spacing approaches the equilibrium spacing of diamond crystal, this band splits into two bands separated by an energy gap $E_g$, where no allowed energy states for electrons exist $\Rightarrow$ forbidden gap.
- The upper band (called the conduction band) and the lower band (called the valence band) contain $4N$ states each.
- The lower 1s band is filled with $2N$ electrons, however, the $4N$ electrons residing in the original $n = 2$ state will now occupy states either in the valence band or in the conduction band.
- At 0 K, the electrons will occupy the lowest energy states available to them $\Rightarrow$ thus, the $4N$ states in the valence band will be completely filled, and the $4N$ states in the conduction band will be completely empty.

Metals, Semiconductors, and Insulators

- For electrons to move under an applied electric field, there must be states available to them.
- A completely filled band cannot contribute to current transport; neither can a completely empty band.
- Thus, semiconductors at 0 K are perfect insulators.
- With thermal or optical excitation, some of these electrons can be excited from the valence band to the conduction band, and then they can contribute to the current transport process.
At temperatures other than 0 K, the magnitude of the band gap separates an insulator from a semiconductor, e.g., at 300 K, $E_g$(diamond) = 5 eV (insulator), and $E_g$(Silicon) = 1.12 eV (semiconductor).

Number of electrons available for conduction can be increased greatly in semiconductors by reasonable amount of thermal or optical energy.

In metals, the bands are either partially filled or they overlap => thus, electrons and empty states coexist => great electrical conductivity.

**Direct and Indirect Semiconductors**

In a typical quantitative calculation of band structures, the wave function of a single electron traveling through a perfectly periodic lattice is assumed to be in the form of a plane wave moving in the x-direction (say) with propagation constant $k$, also called a wave vector.

In quantum mechanics, the electron momentum can be given by $p = \hbar k$.

The space dependent wave function for the electron is

$$\psi_k(x) = U(k_x, x)e^{ikx}$$  \hspace{1cm} (2.1)

where the function $U(k_x, x)$ modulates the wave function according to the periodicity of the lattice.

Allowed values of energy, while plotted as a function of $k$, gives the E-k diagram.

Since the periodicity of most lattices is different in various directions, the E-k diagram is a complex surface, which is to be visualized in three dimensions.

Fig.2.2 Direct and indirect transition of electrons from the conduction band to the valence band: (a) direct - with accompanying photon emission, (b) indirect via defect level.

- Direct band gap semiconductor: the minima of the conduction band and the maxima of the valence band occur at the same value of $k$ => an electron making the smallest energy transition from the conduction band to the valence band can do so without a change in $k$ (and, the momentum).
• Indirect band gap semiconductor: the minima of the conduction band and the maxima of the valence band occur for different values of $k$, thus, the smallest energy transition for an electron requires a change in momentum.
• Electron falling from conduction band to an empty state in valence band $\Rightarrow$ recombination.
• Recombination probability for direct band gap semiconductors is much higher than that for indirect band gap semiconductors.
• Direct band gap semiconductors give up the energy released during this transition ($= E_g$) in the form of light $\Rightarrow$ used for optoelectronic applications (e.g., LEDs and LASERs).
• Recombination in indirect band gap semiconductors occurs through some defect states within the band gap, and the energy is released in the form of heat given to the lattice.

**Variation of Energy Bands with Alloy Composition**

• The band structures of III-V ternary and quaternary compounds change as their composition is varied.
• There are three valleys in the conduction band: $\Gamma$ (at $k = 0$), L, and X.
• In GaAs, the $\Gamma$ valley has the minimum energy (direct with $E_g = 1.43$ eV) with very few electrons residing in L and X valleys (except for high field excitations).
• In AlAs, the X valley has minimum energy (indirect with $E_g = 2.16$ eV).

![E-k diagram of (a) GaAs and (b) AlAs](image)

Fig.2.3 The E-k diagram of (a) GaAs and (b) AlAs, showing the three valleys (L, $\Gamma$, and X) in the conduction band.

**Charge Carriers in Semiconductors**
• In a metal, the atoms are imbedded in a "sea" of free electrons, and these electrons can move as a group under the influence of an applied electric field.
• In semiconductors at 0 K, all states in the valence band are full, and all states in the conduction band are empty.
• At T > 0 K, electrons get thermally excited from the valence band to the conduction band, and contribute to the conduction process in the conduction band.
• The empty states left in the valence band can also contribute to current conduction.
• Also, introduction of impurities has an important effect on the availability of the charge carriers.
• Considerable flexibility in controlling the electrical properties of semiconductors.

Electrons and Holes

• For T > 0 K, there would be some electrons in the otherwise empty conduction band, and some empty states in the otherwise filled valence band.
• The empty states in the valence band are referred to as holes.
• If the conduction band electron and the valence band hole are created by thermal excitation of a valence band electron to the conduction band, then they are called electron-hole pair (EHP).
• After excitation to the conduction band, an electron is surrounded by a large number of empty states, e.g., the equilibrium number of EHPs at 300 K in Si is \( \sim 10^{10}/\text{cm}^3 \), whereas the Si atom density is \( \sim 10^{22}/\text{cm}^3 \).

• Thus, the electrons in the conduction band are free to move about via the many available empty states.
• Corresponding problem of charge transport in the valence band is slightly more complex.
• Current transport in the valence band can be accounted for by keeping track of the holes themselves.
• In a filled band, all available energy states are occupied.
• For every electron moving with a given velocity, there is an equal and opposite electron motion somewhere else in the band.
• Under an applied electric field, the net current is zero, since for every electron \( j \) moving with a velocity \( \vec{v}_j \), there is a corresponding electron \( \vec{v}'_j \) moving with a velocity - \( \vec{v}_j \).
• In a unit volume, the current density \( J \) can be given by

\[
J = (-q) \sum_{i=1}^{N} \vec{v}_i = 0 \tag{2.2}
\]

(filled band) (2.2)

where \( N \) is the number of electrons/cm\(^3\) in the band, and \( q \) is the electronic charge.

• Now, if the \( j^{th} \) electron is removed and a hole is created in the valence band, then the net current density
Thus, the current contribution of the empty state (hole), obtained by removing the jth electron, is equivalent to that of a positively charged particle with velocity \( \hat{v}_j \).

Note that actually this transport is accounted for by the motion of the uncompensated electron \( \hat{v}_j \) having a charge of \( q \) and moving with a velocity \( \hat{v}_j \).

Its current contribution \(- q \hat{v}_j\) is equivalent to that of a positively charged particle with velocity \( + \hat{v}_j \).

For simplicity, therefore, the empty states in the valence band are called holes, and they are assigned positive charge and positive mass.

The electron energy increases as one moves up the conduction band, and electrons gravitate downward towards the bottom of the conduction band.

On the other hand, hole energy increases as one moves down the valence band (since holes have positive charges), and holes gravitate upwards towards the top of the valence band.

**Effective Mass**

The "wave-particle" motion of electrons in a lattice is not the same as that for a free electron, because of the interaction with the periodic potential of the lattice.

To still be able to treat these particles as "free", the rest mass has to be altered to take into account the influence of the lattice.

The calculation of effective mass takes into account the shape of the energy bands in three-dimensional k-space, taking appropriate averages over the various energy bands.

The effective mass of an electron in a band with a given \((E,k)\) relation is given by

\[
m^* = \frac{\hbar^2}{d^2E/dk^2}
\]

**EXAMPLE 2.1:** Find the dispersion relation for a free electron, and, thus, observe the relation between its rest mass and effective mass.

**SOLUTION:** For a free electron, the electron momentum is \( \mathbf{p} = m_0 \mathbf{v} = \hbar \mathbf{k} \). Thus,

\[
E = m_0 \mathbf{v}^2/2 = \frac{\mathbf{p}^2}{2m_0} = \frac{\hbar^2 \mathbf{k}^2}{2m_0}.
\]

Therefore, the dispersion relation, i.e., the E-k relation is parabolic. Hence, \( m^* = \frac{\hbar^2}{\partial^2E/\partial k^2} = m_0 \). This is a very interesting relation, which states that for a free electron, the rest mass and the effective mass are one and the same, which is due to the parabolic band structure. Most materials have non-parabolic E-k relation, and, thus, they have quite different rest mass and effective mass for electrons.

Note: for severely non-parabolic band structures, the effective mass may become a function of...
energy, however, near the minima of the conduction band and towards the maxima of the valence band, the band structure can be taken to be parabolic, and, thus, an effective mass, which is independent of energy, may be obtained.

- Thus, the effective mass is an inverse function of the curvature of the E-k diagram: weak curvature gives large mass, and strong curvature gives small mass.
- Note that in general, the effective mass is a tensor quantity, however, for parabolic bands, it is a constant.
- Another interesting feature is that the curvature $\frac{d^2E}{dk^2}$ is positive at the conduction band minima, however, it is negative at the valence band maxima.
- Thus, the electrons near the top of the valence band have negative effective mass.
- Valence band electrons with negative charge and negative mass move in an electric field in the same direction as holes with positive charge and positive mass.
- Thus, the charge transport in the valence band can be fully accounted for by considering hole motion alone.
- The electron and hole effective masses are denoted by $m^*_e$ and $m^*_p$ respectively.

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>The effective mass for electrons and holes. ($\propto m_0$, where $m_0$ is the rest mass for electrons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m^*_e$</td>
<td>Ge</td>
</tr>
<tr>
<td>0.55</td>
<td>1.1</td>
</tr>
<tr>
<td>$m^*_p$</td>
<td>0.37</td>
</tr>
</tbody>
</table>

**Intrinsic Material**

- A perfect semiconductor crystal with no impurities or lattice defects.
- No carriers at 0 K, since the valence band is completely full and the conduction band is completely empty.
- For $T > 0$ K, electrons are thermally excited from the valence band to the conduction band (EHP generation).
- EHP generation takes place due to breaking of covalent bonds $\Rightarrow$ required energy $= E_g$.
- The excited electron becomes free and leaves behind an empty state (hole).
- Since these carriers are created in pairs, the electron concentration ($n/cm^3$) is always equal to the hole concentration ($p/cm^3$), and each of these is commonly referred to as the **intrinsic carrier concentration** ($n_i$).
- Thus, for intrinsic material $n = p = n_i$.
- These carriers are not localized in the lattice; instead they spread out over several lattice spacings, and are given by quantum mechanical probability distributions.
- Note: $n_i = f(T)$.
- To maintain a steady-state carrier concentration, the carriers must also recombine at the same rate at which they are generated.
• Recombination occurs when an electron from the conduction band makes a transition (direct or indirect) to an empty state in the valence band, thus annihilating the pair.
• At equilibrium, \( \dot{n} = \dot{g} \), where \( \dot{g} \) and \( \dot{n} \) are the generation and recombination rates respectively, and both of these are temperature dependent.
• \( \dot{g}(T) \) increases with temperature, and a new carrier concentration \( n_i \) is established, such that the higher recombination rate \( \dot{n}(T) \) just balances generation.
• At any temperature, the rate of recombination is proportional to the equilibrium concentration of electrons and holes, and can be given by \( \dot{n} = \alpha n_h p_0 = \alpha n_i^2 = g(T) \) (2.5) where \( \alpha \) is a constant of proportionality (depends on the mechanism by which recombination takes place).

Extrinsic Material

• In addition to thermally generated carriers, it is possible to create carriers in the semiconductor by purposely introducing impurities into the crystal \( \Rightarrow \text{doping} \).
• Most common technique for varying the conductivity of semiconductors.
• By doping, the crystal can be made to have predominantly electrons (n-type) or holes (p-type).
• When a crystal is doped such that the equilibrium concentrations of electrons (\( n_0 \)) and holes (\( p_0 \)) are different from the intrinsic carrier concentration (\( n_i \)), the material is said to be extrinsic.
• Doping creates additional levels within the band gap.
• In Si, column V elements of the periodic table (e.g., P, As, Sb) introduce energy levels very near (typically 0.03-0.06 eV) the conduction band.
• At 0 K, these levels are filled with electrons, and very little thermal energy (50 K to 100 K) is required for these electrons to get excited to the conduction band.
• Since these levels donate electrons to the conduction band, they are referred to as the donor levels.
• Thus, Si doped with donor impurities can have a significant number of electrons in the conduction band even when the temperature is not sufficiently high enough for the intrinsic carriers to dominate, i.e., \( n_0 >> n_i \), \( p_0 => n \)-type material, with electrons as majority carriers and holes as minority carriers.
• In Si, column III elements of the periodic table (e.g., B, Al, Ga, In) introduce energy levels very near (typically 0.03-0.06 eV) the valence band.
• At 0 K, these levels are empty, and very little thermal energy (50 K to 100 K) is required for electrons in the valence band to get excited to these levels, and leave behind holes in the valence band.
• Since these levels accept electrons from the valence band, they are referred to as the acceptor levels.
• Thus, Si doped with acceptor impurities can have a significant number of holes in the valence band even at a very low temperature, i.e., \( p_0 >> n_i \), \( n_0 => p \)-type material, with holes as majority carriers and electrons as minority carriers.
• The extra electron for column V elements is loosely bound and it can be liberated very easily =>ionization; thus, it is free to participate in current conduction.
• Similarly, column III elements create holes in the valence band, and they can also participate in current conduction.
• Rough calculation of the ionization energy can be made based on the Bohr's model for $H_2$ atoms, considering the loosely bound electron orbiting around the tightly bound core electrons. Thus,

$$ E = \frac{m_h^* q^4}{2(4\pi \varepsilon_0 \varepsilon_r)^2 \hbar^2} $$

(2.6)where $\varepsilon_r$ is the relative permittivity of Si.

EXAMPLE 2.2: Calculate the approximate donor binding energy for Si ($r = 11.7, m_h^* = 1.18 m_0$).

SOLUTION: From Eq.(2.6), we have

$$ E = \frac{m_h^* q^4}{8(4\pi \varepsilon_0 \varepsilon_r)^2 \hbar^2} = \frac{1.18 \times 9.11 \times 10^{-21} \times 1.6 \times 10^{-19} \ h}{8 \times 8.854 \times 10^{-12} \times 11.7 \ h^2 \times 6.63 \times 10^{-34} \ h^2} = 1.867 \times 10^{-20} \ J = 0.117 \ eV. $$

Note: The effective mass used here is an average of the effective mass in different crystallographic directions, and is called the "conductivity effective mass" with values of 1.28 $m_0$ (at 600 K), 1.18 $m_0$ (at 300 K), 1.08 $m_0$ (at 77 K), and 1.026 $m_0$ (at 4.2 K).

• In III-V compounds, column VI impurities (e.g., S, Se, Te) occupying column V sites act as donors. Similarly, column II impurities (e.g., Be, Zn, Cd) occupying column III sites act as acceptors.

• When a column IV material (e.g., Si, Ge) is used to dope III-V compounds, then they may substitute column III elements (and act as donors), or substitute column V elements (and act as acceptors) => amphoteric dopants.

• Doping creates a large change in the electrical conductivity, e.g., with a doping of $10^{16}/cm^3$, the resistivity of Si changes from $2 \times 10^{6} \ \Omega \cdot cm$ to $5 \ \Omega \cdot cm$.

**Carrier Concentrations**

• For the calculation of semiconductor electrical properties and analyzing device behavior, it is necessary to know the number of charge carriers/cm3 in the material.

• The majority carrier concentration in a heavily doped material is obvious, since for each impurity atom, one majority carrier is obtained.

• However, the minority carrier concentration and the dependence of carrier concentrations on temperature are not obvious.

• To obtain the carrier concentrations, their distribution over the available energy states is required.

• These distributions are calculated using statistical methods.

**The Fermi Level**
Electrons in solids obey *Fermi-Dirac* (FD) statistics.

This statistics accounts for the indistinguishability of the electrons, their wave nature, and the Pauli exclusion principle.

The Fermi-Dirac distribution function \( f(E) \) of electrons over a range of allowed energy levels at thermal equilibrium can be given by

\[
f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}
\]

where \( k \) is Boltzmann's constant (= \( 8.62 \times 10^{-5} \text{eV/K} = 1.38 \times 10^{-23} \text{J/K} \)).

This gives the probability that an available energy state at \( E \) will be occupied by an electron at an absolute temperature \( T \).

\( E_F \) is called the Fermi level and is a measure of the average energy of the electrons in the lattice \( \Rightarrow \) an extremely important quantity for analysis of device behavior.

Note: for \( (E - E_F) > 3kT \) (known as *Boltzmann approximation*), \( f(E) \approx \exp[-(E - E_F)/kT] \Rightarrow \) this is referred to as the Maxwell-Boltzmann (MB) distribution (followed by gas atoms).

The probability that an energy state at \( E_F \) will be occupied by an electron is 1/2 at all temperatures.

At 0 K, the distribution takes a simple rectangular form, with all states below \( E_F \) occupied, and all states above \( E_F \) empty.

At \( T > 0 \) K, there is a finite probability of states above \( E_F \) to be occupied and states below \( E_F \) to be empty.

The F-D distribution function is highly symmetric, i.e., the probability \( f(E_F + \Delta E) \) that a state \( E \) above \( E_F \) is filled is the same as the probability \( [1 - f(E_F - \Delta E)] \) that a state \( E \) below \( E_F \) is empty.

This symmetry about \( E_F \) makes the Fermi level a natural reference point for the calculation of electron and hole concentrations in the semiconductor.

Note: \( f(E) \) is the probability of occupancy of an available state at energy \( E \), thus, if there is no available state at \( E \) (e.g., within the band gap of a semiconductor), there is no possibility of finding an electron there.

For intrinsic materials, the Fermi level lies close to the middle of the band gap (the difference between the effective masses of electrons and holes accounts for this small deviation from the mid gap).

In n-type material, the electrons in the conduction band outnumber the holes in the valence band, thus, the Fermi level lies closer to the conduction band.

Similarly, in p-type material, the holes in the valence band outnumber the electrons in the conduction band, thus, the Fermi level lies closer to the valence band.

The probability of occupation \( f(E) \) in the conduction band and the probability of vacancy \([1 - f(E)]\) in the valence band are quite small, however, the densities of available states in these bands are very large, thus a small change in \( f(E) \) can cause large changes in the carrier concentrations.
Fig. 2.4 The density of states $N(E)$, the Fermi-Dirac distribution function $f(E)$, and the carrier concentration as functions of energy for (a) intrinsic, (b) n-type, and (c) p-type semiconductors at thermal equilibrium.

- Note: since the function $f(E)$ is symmetrical about $E_F$, a large electron concentration implies a small hole concentration, and vice versa.
- In n-type material, the electron concentration in the conduction band increases as $E_F$ moves closer to $E_C$; thus, $(E_C - E_F)$ gives a measure of $n$.
- Similarly, in p-type material, the hole concentration in the valence band increases as $E_F$ moves closer to $E_V$; thus, $(E_V - E_F)$ gives a measure of $p$.

**Electron and Hole Concentrations at Equilibrium**

- The F-D distribution function can be used to calculate the electron and hole concentrations in semiconductors, if the densities of available states in the conduction and valence bands are known.
In equilibrium, the concentration of electrons in the conduction band can be given by

\[ n_0 = \int_{E_c}^{\infty} f(E)N(E)dE \tag{2.8} \]

where \( N(E)dE \) is the density of available states/cm\(^3\) in the energy range \( dE \).

Note: the upper limit of \( E \) is theoretically not proper, since the conduction band does not extend to infinite energies; however, since \( f(E) \) decreases rapidly with increasing \( E \), the contribution to this integral for higher energies is negligible.

Using the solution of Schrödinger's wave equation under periodic boundary conditions, it can be shown that

\[ N(E) = \frac{1}{2\pi^2} \left[ \frac{2m^*}{\hbar^2/2} \right]^{3/2} E^{1/2} \tag{2.9} \]

Thus, \( N(E) \) increases with \( E \), however, \( f(E) \) decreases rapidly with \( E \), thus, the product \( f(E)N(E) \) decreases rapidly with \( E \), and very few electrons occupy states far above the conduction band edge, i.e., most electrons occupy a narrow energy band near the conduction band edge.

Similarly, the probability of finding an empty state in the valence band \([1 - f(E)]\) decreases rapidly below \( E_V \), and most holes occupy states near the top of the valence band.

Thus, a mathematical simplification can be made assuming that all available states in the conduction band can be represented by an effective density of states \( N_C \) located at the conduction band edge \( E_C \) and using Boltzmann approximation.

Thus,

\[ n_0 = N_C f(E_C) = N_C e^{-(E_C - E_F)/kT} \tag{2.10} \]

\[ N_C = \frac{2 \pi m^* n_h kT}{\hbar^2} \left( \frac{2}{3} \right)^{3/2} \]

where \( E_C \).

Note: as \((E_C - E_F)\) decreases, i.e., the Fermi level moves closer to the conduction band, the electron concentration increases.

By similar arguments,

\[ p_0 = N_V [1 - f(E_V)] = N_V e^{-(E_V - E_F)/kT} \tag{2.11} \]

where \( N_V \) is the effective density of states located at the valence band edge \( E_V \).

Note: the only terms separating the expressions for \( N_C \) and \( N_V \) are the effective masses of electrons \((m^*_e)\) and holes \((m^*_h)\) respectively, and since \( m^*_e \), hence, \( N_C \), \( N_V \).
Thus, as \((E_F - E_V)\) decreases, i.e., the Fermi level moves closer to the valence band edge, and the hole concentration increases.

- These equations for \(n_0\) and \(p_0\) are valid in equilibrium, irrespective of the material being intrinsic or doped.
- For intrinsic material \(E_F\) lies at an intrinsic level \(E_i\) (very near the middle of the band gap), and the intrinsic electron and hole concentrations are given by

\[
\begin{align*}
n_i &= N_C e^{-(E_C - E_i)/kT} \\
p_i &= N_V e^{-(E_V - E_i)/kT}
\end{align*}
\]  

and (2.12)

- Note: At equilibrium, the product \(n_0 p_0\) is a constant for a particular material and temperature, even though the doping is varied,

\[
n_0 p_0 = n_i^2 = N_C N_V e^{-E_V/kT}
\]  

(2.13)

- This equation gives an expression for the intrinsic carrier concentration \(n_i\) as a function of \(N_C\), \(N_V\), and temperature:

\[
n_i = \sqrt{N_C N_V e^{-E_V/kT}}
\]  

(2.14)

- These relations are extremely important, and are frequently used for calculations.
- Note: if \(N_C\) were to be equal to \(N_V\), then \(E_i\) would have been exactly at mid gap (i.e., \(E_C - E_F = E_i = E_V = E_a/2\)).
- However, since \(N_C \gg N_V\), \(E_i\) is displaced slightly from mid gap (more for GaAs than that for Si).
- Alternate expressions for \(n_0\) and \(p_0\):

\[
\begin{align*}
n_0 &= n_i e^{(E_F - E_i)/kT} \\
p_0 &= n_i e^{(E_F - E_i)/kT}
\end{align*}
\]  

and (2.15)

- Note: the electron concentration is equal to \(n_i\) when \(E_F\) is at \(E_i\), and \(n_0\) increases exponentially as \(E_F\) moves away from \(E_i\) towards the conduction band.
- Similarly, the hole concentration \(p_0\) varies from \(n_i\) to larger values as \(E_F\) moves from \(E_i\) towards the valence band.

**EXAMPLE 2.3**: A Si sample is doped with \(10^{16}\) B atoms/cm\(^2\). What is the equilibrium electron concentration \(n_0\) at 300 K? Where is \(E_F\) relative to \(E_i\)? Assume \(n_i\) for Si at 300 K = 1.5 \times \(10^{10}\)/cm\(^3\).

**SOLUTION**: Since B (trivalent) is a p-type dopant in Si, hence, the material will be predominantly p-type, and since \(N_A \gg n_i\), therefore, \(p_0\) will be approximately equal to \(N_A\), and

\[
n_0 = \frac{n_i^2}{p_0} = 2.25 \times 10^{20}/10^{16} = 2.25 \times 10^4 \text{ cm}^{-3}
\]

Also,
$E_i - E_F = kT \ln(\rho_0/n_i) = 0.026 \ln[10^{10}/(1.5 \times 10^{10})] = 0.35 \text{ eV}$. The resulting band diagram is:

Temperature Dependence of Carrier Concentrations

- The intrinsic carrier concentration has a strong temperature dependence, given by

  \[
  n_i(T) = 2 \left( \frac{2\pi kT}{\hbar^2} \right)^{3/2} (m_e^* m_b^*)^{3/4} e^{-E_g/2kT}
  \]  

  (2.16)

- Thus, explicitly, $n_i$ is proportional to $T^{3/2}$ and to $e^{-1/T}$, however, $E_g$ also has a temperature dependence (decreasing with increasing temperature, since the interatomic spacing changes with temperature).

Fig. 2.5 The intrinsic carrier concentration as a function of inverse temperature for Si, Ge, and GaAs.
As \( n \) changes with temperature, so do \( \rho \) and \( \beta \).

With \( n \) and \( T \) given, the unknowns are the carrier concentrations and the Fermi level position with respect to \( E_F \); one of these quantities must be given in order to calculate the other.

Example: Si doped with \( 10^{15}/\text{cm}^3 \) donors (\( N_d \)).

At very low temperature, negligible intrinsic EHPs exist, and all the donor electrons are bound to the donor atoms.

As temperature is raised, these electrons are gradually donated to the conduction band, and at about 100 K (1000/T = 10), almost all these electrons are donated \( \Rightarrow \) this temperature range is called the ionization region.

Once all the donor atoms are ionized, the electron concentration \( n_0 \approx N_d \), since for each donor atom, one electron is obtained.

![Variation of carrier concentration with inverse temperature](image)

Fig.2.6 Variation of carrier concentration with inverse temperature clearly showing the three regions: ionization, extrinsic, and intrinsic.

Thus, \( n_0 \) remains virtually constant with temperature for a wide range of temperature (called the extrinsic region), until the intrinsic carrier concentration \( n_i \) starts to become comparable to \( N_d \).

For high temperatures, \( n_i \gg N_d \), and the material loses its extrinsic property (called the intrinsic region).

Note: in the intrinsic region, the device loses its usefulness \( \Rightarrow \) determines the maximum operable temperature range.

**Compensation and Space Charge Neutrality**

- Semiconductors can be doped with both donors (\( N_d \)) and acceptors (\( N_A \)) simultaneously.
• Assume a material doped with \( N_d > N_A \) \( \Rightarrow \) predominantly n-type \( \Rightarrow \) \( E_F \) lies above \( E_i \) \( \Rightarrow \) acceptor level \( E_A \) completely full, however, with \( E_F \) above \( E_i \), the hole concentration cannot be equal to \( N_A \).
  
  • Mechanism:
    - Electrons are donated to the conduction band from the donor level \( E_d \).
    - An acceptor state gets filled by a valence band electron, thus creating a hole in the valence band.
    - An electron from the conduction band recombines with this hole.
    - Extending this logic, it is expected that the resultant concentration of electrons in the conduction band would be \( N_d \) \( N_A \) instead of \( N_d \).
    - This process is called compensation.
  
  • By compensation, an n-type material can be made intrinsic (by making \( N_A = N_d \)) or even p-type (for \( N_A > N_d \)).

  Note: a semiconductor is neutral to start with, and, even after doping, it remains neutral (since for all donated electrons, there are positively charged ions \( N_d^+ \); and for all accepted electrons (or holes in the valence band), there are negatively charged ions \( N_A^- \)).

  Therefore, the sum of positive charges must equal the sum of negative charges, and this governing relation, given by \( p_0 + N_A^- = n_0 + N_d^+ \) (2.17) is referred to as the equation for space charge neutrality.

  This equation, solved simultaneously with the law of mass action (given by \( n_0 p_0 = n_i^2 \)) gives the information about the carrier concentrations.

  Note: for \( \frac{N_d^+ - N_A^-}{N_i^2} \gg \frac{n_i}{n_0} \), \( n_0 \approx N_d^+ - N_A^- \), and, similarly, for \( \frac{N_A^- - N_d^+}{p_0} \gg \frac{n_i^2}{p_0} \), \( p_0 \approx N_d^+ - N_A^- \).

Drift of Carriers in Electric and Magnetic Fields

• In addition to the knowledge of carrier concentrations, the collisions of the charge carriers with the lattice and with the impurity atoms (or ions) under electric and/or magnetic fields must be accounted for, in order to compute the current flow through the device.

• These processes will affect the ease (mobility) with which carriers move within a lattice.

• These collision and scattering processes depend on temperature, which affects the thermal motion of the lattice atoms and the velocity of the carriers.

Conductivity and Mobility

• Even at thermal equilibrium, the carriers are in a constant motion within the lattice.

• At room temperature, the thermal motion of an individual electron may be visualized as random scattering from lattice atoms, impurities, other electrons, and defects.

• There is no net motion of the group of \( \text{electrons/cm}^3 \) over any period of time, since the scattering is random, and there is no preferred direction of motion for the group of electrons and no net current flow.

• However, for an individual electron, this is not true the probability of an electron returning to its starting point after time \( t \) is negligibly small.
Now, if an electric field $E_x$ is applied in the x-direction, each electron experiences a net force $qE_x$ from the field.

This will create a net motion of group in the x-direction, even though the force may be insufficient to appreciably alter the random path of an individual electron.

If $p_x$ is the x-component of the total momentum of the group, then the force of the field on the $n$ electrons/cm$^3$ is

$$-nqE_x = \left. \frac{dp_x}{dt} \right|_{\text{field}}$$  (2.18)

**Note:** This expression indicates a constant acceleration in the x-direction, which realistically cannot happen.

In steady state, this acceleration is just balanced by the deceleration due to the collisions.

Thus, while the steady field $E_x$ does produce a net momentum $p_{x,0}$ for steady state current flow, the net rate of change of momentum must be zero when collisions are included.

Note: the collision processes are totally random, thus, there is a constant probability of collision at any time for each electron.

Consider a group of $N_0$ electrons at time $t = 0$, and define $N(t)$ as the number of electrons that have not undergone a collision by time $t$.

![The random thermal motion of an individual electron, undergoing random scattering.](image)

The rate of decrease of $N(t)$ at any time $t$ is proportional to the number left unscattered at $t$, i.e.

$$- \frac{dN(t)}{dt} = \frac{1}{\bar{\tau}} N(t)$$  (2.19)

where $\bar{\tau}$ is the constant of proportionality.

The solution is an exponential function

$$N(t) = N_0 e^{-t/\bar{\tau}}$$  (2.20)

and $\bar{\tau}$ represents the mean time between scattering events, called the mean free time.
The probability that any electron has a collision in time interval \( dt \) is \( \frac{dt}{\bar{t}} \), thus, the differential change in \( p_x \) due to collisions in time \( dt \) is

\[
\frac{dp_x}{dt} = -p_x \frac{dt}{\bar{t}} \quad (2.21)
\]

Thus, the rate of change of \( p_x \) due to the decelerating effect of collisions is

\[
\frac{dp_x}{dt} \bigg|_{\text{collisions}} = -\frac{p_x}{\bar{t}} \quad (2.22)
\]

For steady state, the sum of acceleration and deceleration effects must be zero, thus,

\[
-\frac{p_x}{\bar{t}} - n q E_x = 0
\]

(2.23)

The average momentum per electron (averaged over the entire group of electrons) is

\[
< p_x > = \frac{p_x}{n} = -q \bar{t} \frac{E_x}{m_n}
\]

(2.24)

Thus, as expected for steady state, the electrons would have on the average a constant net velocity in the -x-direction

\[
< v_x > = \frac{< p_x >}{m_n} = -\frac{q \bar{t} E_x}{m_n}
\]

(2.25)

This speed is referred to as the drift speed, and, in general, it is usually much smaller than the random speed due to thermal motion \( v_{th} = \sqrt{\frac{2kT}{m_n}} \).

The current density resulting from this drift

\[
J_x = -qn < v_x > = \frac{nq^2 \bar{t}}{m_n} E_x = \sigma E_x
\]

(2.26)

This is the familiar Ohm's law with \( \sigma \) being the conductivity of the sample, which can also be written as \( \sigma = q \mu_n \), with \( \mu_n = \frac{q \bar{t}}{m_n} \) being the electron mobility (in \( \text{cm}^2/\text{V-sec} \)), and it describes the ease with which electrons drift in the material.

The mobility can also be expressed as the average drift velocity per unit electric field, thus \( \mu_n = -< v_x > / E_x \) with the negative sign denoting a positive value for mobility since electrons drift opposite to the direction of the electric field.

The total current density can be given by

\[
J_x = q(\mu_n + \mu_P) E_x
\]

(2.27) when both electrons and holes contribute to the current conduction; on the other hand, for predominantly n-type or p-type samples, respectively the first or the second term of the above equation dominates.
Note: both electron and hole drift currents are in the same direction, since holes (with positive charges) move along the direction of the electric field, and electrons (with negative charges) drift opposite to the direction of the electric field.

- Since GaAs has a strong curvature of the E-k diagram at the bottom of the conduction band, the electron effective mass in GaAs is very small \( \Rightarrow \) the electron mobility in GaAs is very high since \( \mu \) is inversely proportional to \( m^* \).

- The other parameter in the mobility expression, i.e., \( \bar{\tau} \) (the mean free time between collisions) is a function of temperature and the impurity concentration in the semiconductor.

- For a uniformly doped semiconductor bar of length L, width w, and thickness t, the resistance \( R \) of the bar can be given by \( \frac{L}{(wt\sigma)} = \rho \frac{L}{(wt)} \) where \( \rho \) is the resistivity.

Effects of Temperature and Doping on Mobility

- The two main scattering events that influence electron and hole motion (and, thus, mobility) are the lattice scattering and the impurity scattering.
- All lattice atoms vibrate due to temperature and can scatter carriers due to collisions.
- These collective vibrations are called phonons, thus lattice scattering is also known as phonon scattering.
- With increasing temperature, lattice vibrations increase, and the mean free time between collisions decreases \( \Rightarrow \) mobility decreases (typical dependence \( \sim T^{-3/2} \)).
- Scattering from crystal defects and ionized impurities dominate at low temperatures.
- Since carriers moving with low velocity (at low temperature) can get scattered more easily by ionized impurities, this kind of scattering causes a decrease in carrier mobility with decreasing temperature (typical dependence \( \sim T^{3/2} \)).
- Note: the scattering probability is inversely proportional to the mean free time (and to mobility), hence, the mobilities due to two or more scattering events add inversely: \( \frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \ldots \) (2.28)
- Thus, the mechanism causing the lowest mobility value dominates.
- Mobility also decreases with increasing doping, since the ionized impurities scatter carriers more (e.g., \( \mu \) for intrinsic Si is 1350 cm\(^2\)/V-sec at 300 K, whereas with a donor doping of \( 10^{17} \) cm\(^-3\), n drops to 700 cm\(^2\)/V-sec).

High Field Effects

- For small electric fields, the drift current increases linearly with the electric field, since \( \sigma \) is a constant.
- However, for large electric fields (typically > \( 10^3 \) V/cm), the current starts to show a sublinear dependence on the electric field and eventually saturates for very high fields.
- Thus, becomes a function of the electric field, and this is known as the hot carrier effect, when the carrier drift velocity becomes comparable to its thermal velocity.
• The maximum carrier drift velocity is limited to its mean thermal velocity (typically $10^7 \text{ cm/sec}$), beyond which the added energy imparted by the electric field is absorbed by the lattice (thus generating heat) instead of a corresponding increase in the drift velocity.

2.4.4 The Hall Effect

• An extremely important measurement procedure for determining the majority carrier concentration and mobility.

![Fig.2.8 The experimental setup for the Hall Effect measurement.](image)

- If a magnetic field is applied perpendicular to the direction of carrier flow, the path of the carriers get deflected due to the Lorentz force experienced by the carriers, which can be given by
  \[ F = q(E + v \times B) \]  
- Thus, the holes will get deflected towards the -y-direction, and establish an electric field along the y-direction, such that in steady state \[ E_y = v_B B_z \]
- The establishment of this electric field is known as the Hall effect, and the resulting voltage \[ V_{AB} = E_y w \] is called the Hall voltage.
- Using the expression for the drift current, \[ E_y = R_H J_z B_z, \] where \[ R_H = \frac{1}{\sigma p_0} \] is called the Hall coefficient.
- A measurement of the Hall voltage along with the information for magnetic field and current density gives the majority carrier concentration \[ p_0 = \frac{I_z B_z}{[\sigma V_{AB}]} \].
- Also, the majority carrier mobility \[ \mu_e = \frac{R_H}{\rho} \] can be obtained from a measurement of the resistivity \[ \rho = \frac{[\sigma V_{CB} L]}{[L(\omega t)]} \].
- This experiment can be performed to obtain the variation of majority carrier concentration and mobility as a function of temperature.
- For n-type samples, the Hall voltage and the Hall coefficient are negative \( \Rightarrow \) a common diagnostic tool for obtaining the sample type.
- Note: caution should be exercised for near intrinsic samples.
EXAMPLE 2.4: A sample of Si is doped with $10^{18}$ In atoms/cm$^3$. What will be the measured value of its resistivity? What is the expected Hall voltage in a 150 $\mu$m thick sample if $I_x = 2$ mA and $B_z = 5$ kG ($1 \text{kG} = 10^{-5}$ Wb/cm$^2$)?

SOLUTION:

The sample is p-type with $p_0 >> n_0$, and the mobility $\mu_p \approx 400 \text{cm}^2/V\text{-sec}$. Thus, the resistivity $\rho = (\sigma)^{-1} = (q\mu_p p_0)^{-1} = 1.5625 \Omega\text{-cm}$.

The Hall coefficient $R_H = + (q\mu_p)^{-1} = +625 \text{cm}^3/C$ (note the unit of $R_H$).

Note, alternately $R_H = \rho_{\mu_p} = 625 \text{cm}^3/C$.

The Hall voltage $V_{AB} = (I_x B_z R_H)/t = (2 \times 10^{-3})(5 \times 10^{-5})(625)/(150 \times 10^{-4}) = 4.17 \text{mV}$.

Equilibrium Condition

- In equilibrium, there is no external excitation except a constant temperature, no net transfer of energy, no net carrier motion, and no net current transport.

- An important condition for equilibrium is that no discontinuity or gradient can arise in the equilibrium Fermi level $E_F$.

- Assume two materials 1 and 2 (e.g., n- and p-type regions, dissimilar semiconductors, metal and semiconductor, two adjacent regions in a nonuniformly doped semiconductor) in intimate contact such that electron can move between them.

- Assume materials 1 and 2 have densities of state $N_1(E)$ and $N_2(E)$, and F-D distribution functions $f_1(E)$ and $f_2(E)$ respectively at any energy $E$.

  The sample is p-type with $p_0 >> n_0$, and the mobility $\mu_p \approx 400 \text{cm}^2/V\text{-sec}$. Thus, the resistivity $\rho = (\sigma)^{-1} = (q\mu_p p_0)^{-1} = 1.5625 \Omega\text{-cm}$.

  The Hall coefficient $R_H = + (q\mu_p)^{-1} = +625 \text{cm}^3/C$ (note the unit of $R_H$).

  Note, alternately $R_H = \rho_{\mu_p} = 625 \text{cm}^3/C$.

  The Hall voltage $V_{AB} = (I_x B_z R_H)/t = (2 \times 10^{-3})(5 \times 10^{-5})(625)/(150 \times 10^{-4}) = 4.17 \text{mV}$.

- The rate of electron motion from 1 to 2 can be given by rate from 1 to 2 $N_1(E)f_1(E)$. $N_2(E)[1 f_2(E)]$ (2.30) and the rate of electron motion from 2 to 1 can be given by rate from 2 to 1 $N_2(E)f_2(E)$. $N_1(E)[1 f_1(E)]$ (2.31). At equilibrium, these two rates must be equal, which gives $f_1(E) = f_2(E)$, $E_F1 = E_F2$ $\Rightarrow$ $dE_F/dx = 0$; thus, the Fermi level is constant at equilibrium, or, in other words, there cannot be any discontinuity or gradient in the Fermi level at equilibrium.
Practice Problems

2.1 Electrons move in a crystal as wave packets with a group velocity \( v = \frac{d\omega}{dk} \), where \( \omega \) is the angular frequency. Show that in a given electric field, these wave packets obey Newton’s second law of motion, i.e., the force \( F = m^*a \), where \( m^* \) is the effective mass (given by \( m^* = \frac{\hbar^2}{2\beta \frac{\partial E}{\partial \mathbf{k}^2}} \)) and \( a \) is the acceleration.

2.2 Some semiconductors of interest have the dependence of its energy \( E \) with respect to the wave vector \( k \), given by \( E(1 + \alpha E) = \frac{\hbar^2 k^2}{2m^*} \), where \( m^* \) is the effective mass for \( E = 0 \), \( k \) is the wave vector, and \( \alpha \) is a constant. Calculate the dependence of the effective mass \( m^* \) on energy.

2.3 Determine the equilibrium recombination constant \( r \) for Si and GaAs, having equilibrium thermal generation rates of \( 4.03 \times 10^6 \text{ cm}^3\text{-sec}^{-1} \) and \( 3.18 \times 10^8 \text{ cm}^3\text{-sec}^{-1} \) respectively, and intrinsic carrier concentrations of \( 1.5 \times 10^{10} \text{ cm}^{-3} \) and \( 2.1 \times 10^8 \text{ cm}^{-3} \) respectively. Comment on the answers. Will \( \alpha \) change with doping at equilibrium?

2.4 The relative dielectric constant for GaP is 10.2 and the electron effective mass is \( 0.13m_0 \). Calculate the approximate ionization energy of a donor atom in GaP.

2.5 Show that the probability that a state \( \Delta E \) above the Fermi level \( E_F \) is occupied is the same as the probability that a state \( \Delta E \) below \( E_F \) is empty.

2.6 Derive an expression relating the intrinsic level \( E_i \) to the center of the band gap \( E_g/2 \), and compute the magnitude of this displacement for Si and GaAs at 300 K. Assume \( m^*_n/m_0 = 1.1 \) and 0.067, and \( m^*_p/m_0 = 0.56 \) and 0.48 for Si and GaAs respectively.

2.7 Show that in order to obtain maximum resistivity in a GaAs sample \( (E_g = 1.42 \text{ eV}, n_i = 2.1 \times 10^6 \text{ cm}^{-3}, \mu_n = 8500 \text{ cm}^2/\text{V}\cdot\text{sec}, \text{ and } \mu_p = 400 \text{ cm}^2/\text{V}\cdot\text{sec}) \), it has to be doped slightly p-type. Determine this doping concentration. Also, determine the ratio of the maximum resistivity to the intrinsic resistivity.

2.8 A GaAs sample (use the date given in Problem 2.7) is doped uniformly with \( 10^7 \text{ Si atoms/cm}^3 \), out of which 70% occupy Ga sites, and the rest 30% occupy As sites. Assume 100% ionization and \( T = 300 \text{ K} \).
   a) Calculate the equilibrium electron and hole concentrations \( n_0 \) and \( p_0 \).
   b) Clearly draw the equilibrium band diagram, showing the position of the Fermi level \( E_F \) with respect to the intrinsic level \( E_i \), assuming that \( E_i \) lies exactly at midgap.
   c) Calculate the percentage change in conductivity after doping as compared to the intrinsic case.
2.9 A Si sample is doped with $10^{16}$ cm$^{-3}$ donor atoms. Determine the minimum temperature at which the sample becomes intrinsic. Assume that at this minimum temperature, the free electron concentration does not exceed by more than 1% of the donor concentration (beyond its extrinsic value). For Si, $n_i(T) = 3.68 \times 10^{16} T^{3/2} e^{-7000/T}$ cm$^{-3}$ (T in K).

2.10 Since the event of collision of an electron in a lattice is a truly random process, thus having a constant probability of collision at any given time, the number of particles left unscattered at time $t$, $N(t) = N_0 \exp(-t/\tau)$, where $N_0 = N(t)$ at $t = 0$, and $\tau = \text{mean free time}$. Hence, show that if there are a total of $i$ number of scattering events, each with a mean free time of $\frac{1}{\mu}$, then the net electron mobility $\mu$ can be given by $\mu = \sum_{i}^{-1} (\mu_i)^{-1}$, where $\mu_i$ is the mobility due to the $i$th scattering event.

2.11 A Ge sample is oriented in a 5 kG (1 kG = 10$^{-5}$ Wb/cm$^2$) magnetic field (refer to Fig.2.8). The current is 4 mA, and the sample dimensions are $w = 0.25$ mm, $t = 50$ $\mu$m, and $L = 2.5$ mm. The following data are taken: $V_{AB} = -2.5$ mV, and $V_{CD} = 170$ mV. Find the type and concentration of the majority carrier, and its mobility. Hence, compute the net relaxation time for the various scattering events, assuming $m^*_n = 0.55m_0$ and $m^*_p = 0.37m_0$.

2.12 In the Hall effect experiment, there is a chance that the Hall Probes A and B (refer to Fig.2.8) are not perfectly aligned, which may give erroneous Hall voltage readings. Show that the true Hall voltage $V_H$ can be obtained from two measurements of $V_{AB}$, with the magnetic field first in the +z-direction, and then in the z-direction.

EXCESS CARRIERS IN SEMICONDUCTORS

- Excess carriers, essential for device operation, are created by optical excitation, electron bombardment, or injected across a forward-biased p-n junction.
- These excess carriers can dominate the conduction process in semiconductor materials.

Optical Absorption

- This includes photons in the optical range as well as those in the infrared region.
- Photons of various wavelengths (frequencies) are directed at the sample, and their relative transmission is measured.
- Note: photons having energies greater than the band gap energy are absorbed (the sample behaves opaque for this kind of illumination), whereas those having energies less than the band gap energy are transmitted (the sample behaves transparent), this experiment gives an accurate measure of the band gap energy.
• When photons having energies $h\Delta E_g$ are absorbed, they create EHPs and the probability of this absorption is very high, since there are lots of electrons in the valence band and lots of empty states in the conduction band.

• Electrons excited to EC may initially have energies much higher than EC, however, they lose this excess energy due to scattering with the lattice until their equilibrium energy becomes equal to EC.

• Note: these EHPs are called excess carriers, since they are out of balance, and, thus, would eventually recombine.

• However, while these excess carriers remain in the respective bands, they can contribute to the current conduction.

• The transmitted intensity $I_t$ of a beam of photons of wavelength through a sample of thickness $t$ can be given by

$$I_t = I_0 e^{-\alpha t}$$

where $\alpha$ (cm$^{-1}$) is called the absorption coefficient, and varies with materials and photon wavelength.

Fig.3.1 The variation of the absorption coefficient as a function of the wavelength of the incident light.

Fig.3.2 Band gaps of some common semiconductors relative to the optical spectrum.
Absorption cutoff occurs at \( \lambda_c = 1.24/E_g \) (\( E_g \) in eV).

GaAs, Si, Ge, and InSb band gaps are such that \( \lambda_c \) occurs beyond the visible region (in the infrared), whereas GaP and CdS have band gaps with \( \lambda_c \) falling within the visible range.

**Luminescence**

- When recombination occurs between a conduction band electron and a valence band hole, the energy released can be given off in the form of light (luminescence). Direct band-to-band recombination in direct band gap semiconductors have a much higher probability of light emission as compared to those in indirect materials.
- Broadly divided into three categories:
  - Photoluminescence: if the recombining carriers were caused by optical excitation.
  - Cathodoluminescence: if the recombining carriers were caused by high energy electron bombardment.
  - Electroluminescence: if the recombining carriers were caused by injection of excess carriers (by forward biasing a p-n junction, for example).

**Photoluminescence**

- For steady state excitation, the recombination rate and the generation rate for EHPs are equal, and one photon is emitted for each photon absorbed.
- Direct band-to-band recombination is a fast process with typical lifetime of excess carriers 10^8 sec => known as fluorescence (example: fluorescent lamp).
- In some indirect materials, the trap states within the band gap captures carriers, and slows down the recombination process, thus, emission continues for seconds or minutes after the excitation is removed => known as phosphorescence and the materials are known as phosphors.
- The trap states can hold the carriers for indefinite times, and the carriers can either get reexcited to the conduction band or fall to the valence band (and, thus, recombine) => this creates the delay between excitation and recombination.

**EXAMPLE 3.1:** A 0.5 m thick sample of In is illuminated with monochromatic light of \( h\nu = 1.5 \text{ eV} \). The absorption coefficient \( 10^4 \text{ cm}^{-1} \). The power incident on the sample is 15 mW.

(a) Find the total energy absorbed by the sample per second (J/sec).
(b) Find the rate of excess thermal energy given up by the electrons to the lattice before recombination (J/sec).
(c) Find the number of photons per second given off from recombination events, assuming perfect quantum efficiency.
SOLUTION:

a. (a) The transmitted intensity
\[ I = I_0 e^{-15} = 15 \times 10^{-3} \exp(-10^4 \times 0.5 \times 10^{-4}) = 9.1 \text{ mW}. \]

Therefore, the absorbed power
\[ (15 \times 9.1) \text{ mW} = 5.9 \text{ mW} = 5.9 \times 10^3 \text{ J/sec}. \]

b. (b) Since the energy of the incident photon is greater than the band gap energy, hence, the excess energy of the excited electron will be dissipated as heat to the lattice. The fraction of energy converted to heat is given by
\[ \frac{(1.5 \times 1.34)}{1.5} = 0.107. \]
Thus, the amount of energy converted to heat per second
\[ 0.107 \times 5.9 \times 10^{-3} = 0.63 \times 10^{-3} \text{ J/sec}. \]

c. (c) For perfect quantum efficiency, one photon is emitted for each photon absorbed. Thus, the number of photons emitter per second
\[ \frac{5.9 \times 10^{-3} \text{ (J/sec)}}{1.6 \times 10^{-19} \text{ (J/eV)} \times 1.5 \text{ eV/photons)} = 2.46 \times 10^{16} \text{ photons/sec}. \]

or, alternately, recombination radiation accounts for 5.9 \times 0.63 = 5.27 mW at 1.34 eV/photon. Thus,
\[ \frac{5.27 \times 10^{-3}}{1.6 \times 10^{-19} \times 1.34} = 2.46 \times 10^{16} \text{ photons/sec}. \]

---

Cathodoluminescence

- Best example: cathode ray tube (CRT) basis of television sets, oscilloscopes, and other display systems.
- Electrons emitted from the heated cathode are accelerated towards the anode by high field, deflected by electric or magnetic fields by the horizontal and vertical plates, and made to hit the screen (coated with a phosphor) at desired locations.
• When electrons hit these phosphors, the energy of the electrons gets transferred to the electrons of the phosphors, and they get excited to higher states, and eventually fall to the ground state, thus causing recombination and light emission.
• Three phosphor dots are used for each pixel, capable of transmitting three primary colors: red, green, and blue (RGB) thus by varying the intensity and position of the electron beam, a wide range of colors and picture can be attained.

**Electroluminescence**

• Best examples: LEDs and LASERs, where carriers are injected across a forward biased p-n junction and are made to recombine (either naturally or by carrier confinement) => called injection electroluminescence.

**Carrier Lifetime and Photoconductivity**

• Photoconductivity: increase in the conductivity of a sample due to excess carriers created by optical excitation.
• With excitation turned off, the photoconductivity decreases to zero since all excess carriers eventually recombine.

**Direct Recombination of Electrons and Holes**

• Direct recombination occurs spontaneously, i.e., the probability that an electron and a hole will recombine is constant in time, which leads to an exponential solution for the decay of the excess carriers.
• The net rate of change in the conduction band electron concentration at any time $t$

$$\frac{dn(t)}{dt} = \alpha_n n^2 - \alpha_p n(t)p(t) \quad (3.2)$$

where the first term is the generation rate and the second term is the recombination rate.
• Let excess EHPs $n$ and $p$ (with $n = p$, since they are created in pairs) are created at $t = 0$ by a short flash of light.
• Define $n(t)$ and $p(t)$ (again $n = p$) as the instantaneous excess carrier concentrations and $n$ and $p$ for their values at $t = 0$.
• Note: $n(t) = n_0 + n(t)$, and $p(t) = p_0 + p(t)$.
• Thus,
where \( \tau_m \) (\( = [\kappa_r \rho_0]^{-1} \)) is called the minority carrier recombination lifetime or simply the minority carrier lifetime, and it determines the rate at which the minority carriers recombine with time.

- Similarly, excess holes in an n-type material recombine with a rate \( \tau_n \) (\( = [\kappa_f \rho_0]^{-1} \)).
- Note: for direct recombination, the excess majority carriers (which is equal to the excess minority carriers) decay at exactly the same rate as the minority carriers, however, there is a large percentage change in the minority carriers as compared to the majority carriers. 
- A more general expression for carrier lifetime for near not sufficiently extrinsic samples is \( \tau = 1/[\kappa_d(n_0 + \rho_0)] \).

**Indirect Recombination: Trapping**

- In indirect materials, the probability of direct band-to-band recombination is very small recombination in these materials proceed through the assistance of recombination (or trapping) centers located within the band gap, which trap carriers of one type followed by trapping carriers of the opposite type, thus annihilating the pairs.
Fig. 3.3 Energy levels of common impurities in silicon, measured from the closest band edge.

- The resulting energy loss is often in the form of heat given to the lattice (instead of light emission), and, thus, these materials are not well suited for optoelectronic applications.
- There are four probabilities associated with a recombination center:
  (a) hole capture: when an electron from the recombination center falls to the valence band,
  (b) hole emission: when an electron makes a transition from the valence band to the recombination center,
  (c) electron capture: when an electron falls from the conduction band to the recombination center, and
  (d) electron emission: when an electron makes a transition from the recombination center to the conduction band.

- Each of these processes has their own probabilities and time constants, and the resulting analysis is significantly complicated.
- This theory of recombination is known as the SHR (Shockley-Hall-Read) theory of recombination.
- If process (a) follows process (c) or vice versa, recombination takes place, whereas, if process (b) follows process (a) (or vice versa) or process (c) follows process (d) (or vice versa), it is known as reemission, and the recombination center behaves like a trapping center.
Generally, centers that are located towards the middle of the band gap (e.g., Au in Si) behave like recombination centers, whereas centers located closer to the conduction or valence band behave as traps, for obvious reasons.

Alternate definition: in a center located within the band gap, if after capturing one type of carrier, the most probable next event is the capture of opposite type of carrier, then it is a recombination center, however, if the most probable next event is reexcitation, then it is a trap.

The recombination can be slow or fast, depending on the amount of time the carrier spends in the center before the capture of the opposite type of carrier happens, thus, computation of lifetime for this kind of indirect recombination is sufficiently complicated.

The decay of excess carriers can be measured by a typical photoconductive decay measurement, where light shining on a sample is suddenly switched off, and the resulting decay of current passing through the sample is measured, the rate of decay of this current gives the excess carrier lifetime.

**Auger Recombination**

- Note: the lifetime is proportional to the inverse of the doping concentration.
- However, at relatively high doping levels, the lifetime decreases at a faster rate with an increase in the doping concentration.
- This is because a different recombination mechanism, called the Auger recombination becomes dominant at high doping levels.
- In this recombination mechanism, the electron and hole recombine without involving trap levels, and the released energy (of the order of the energy gap) is transferred to another carrier (a hole in p-type material and an electron in n-type material).
- This process is somewhat the inverse of the impact ionization process, in which an energetic carrier causes EHP generation.
- Since two electrons (in n-type material) or two holes (in p-type material) are involved in this process, it is highly unlikely except in heavily doped material.
- The recombination lifetime associated with the Auger recombination process is inversely proportional to the square of the majority carrier concentration, i.e., for p-type material, \( \tau_p \approx 1/(G_p N_A^2) \), and for n-type material, \( \tau_n \approx 1/(G_n N_D^2) \), where \( G_p \) and \( G_n \) are coefficients with values of \( 9.9 \times 10^{-32} \text{ cm}^6/\text{sec} \) and \( 2.28 \times 10^{-31} \text{ cm}^6/\text{sec} \).

**Surface Recombination**

- It is obvious that near the surface of any semiconductor device, the carrier recombination rate should be very high, due to extra defects and traps at the surface.
- Thus, the diffusion flux of minority carriers at the surface is determined by the surface recombination processes.
Fig. 3.4 Auger recombination: (a) n-type sample, and (b) p-type sample.

- For example, when minority carriers are holes, this surface recombination can be described by:

\[
D_p \left. \frac{\partial \rho_n}{\partial x} \right|_{x=0} = -S_p \left[ \rho_n (x=0) - \rho_{n0} \right]
\]  \hspace{1cm} (3.6)

- where \( x = 0 \) corresponds to the surface of the sample, and \( S_p = \sigma_p \gamma_{bo} N_{st} \) is the surface recombination rate, with \( \sigma_p \) being the capture cross-section for holes, \( \gamma_{bo} \) is the thermal velocity for the holes, and \( N_{st} \) is the surface density of the surface states.
- The capture cross-section (typically \( 10^{-15} \text{ cm}^2 \)) describes the effectiveness of the localized state in capturing a carrier.
- The product \( \gamma_{bo} \) may be visualized as the volume swept out per unit time by a particle with cross-section \( \sigma \).
- If the surface state lies within this volume around the carrier, then the carrier gets captured by the surface state.
- Note: the dimension of \( S \) is cm/sec, and, consequently, it is termed as the surface recombination velocity, even though it has nothing to do with actual velocity.
Steady State Carrier Generation: Quasi-Fermi Levels

- For any temperature T, there is a thermal generation rate \( g(T) \) balanced by a recombination rate \( r(T) \).
- Now, if a steady light is shone on the sample, an optical generation rate \( g_{op} \) will be added to \( g(T) \), and the carrier concentrations \( n \) and \( p \) would increase to their new steady state values.
- Generation/recombination rate balance equation:
  \[
g(T) + g_{op} = \alpha_r n p = \alpha_r (n_0 + \delta n)(p_0 + \delta p)
  \]  
  (3.7)
- For steady state recombination and no trapping, \( \delta n = \delta p \); and, under low level injection approximation
  \[
g_{op} = \alpha_r (n_0 + p_0)\delta n = \delta n/\tau_n
  \]  
  (3.8)
- Thus, the excess carrier concentrations can be given by \( \delta n = \delta p = g_{op}\tau_n \).
- In general, when trapping is present, \( \tau_n \neq \tau_p \), and \( \delta n = g_{op}\tau_n \) and \( \delta p = g_{op}\tau_0 \) and \( \delta p = g_{op}\tau_0 \).
- Note: when excess carriers are present, \( np \neq n_1^2 \).
- When excess carriers are present, the equilibrium Fermi level \( E_F \) is no more meaningful; instead, the carrier concentrations are defined in terms of quasi-Fermi levels (also referred to as Imref, which is Fermi spelled backwards) \( F_n \) and \( F_p \) as
  \[
n = n_i e^{(E_n - E_f)/kT} \quad \text{and} \quad p = n_i e^{(E_p - E_f)/kT}
  \]  
  (3.9)
- Imref for the minority carriers deviates significantly from the equilibrium Fermi level, whereas, for majority carriers, the Imref stays very close to the equilibrium Fermi level, and the separation between these two Imrefs is a measure of how far the system is from equilibrium.
- With concentrations varying with position, the Imrefs would also vary with position, thus drawing Imrefs in band diagrams clearly shows the positional variations of the carrier concentrations.

EXAMPLE 3.2

A Si sample with \( p_0 = 10^{15} \text{ cm}^{-3} \) is illuminated by a steady light thus creating \( 10^{12} \text{ EHP/cm}^2 \cdot \mu \text{sec} \) optically. Assume no trapping, and \( \tau_n = \tau_p = 5 \mu \text{sec} \).
(a) Determine the electron and hole concentrations \( n \) and \( p \) respectively, and their percentage change from the equilibrium concentrations.
(b) Comment on the magnitude of the product \( np \).
(c) Determine the locations of the Imrefs \( F_n \) and \( F_p \), and compare their locations with the equilibrium Fermi level \( E_F \).
SOLUTION

(a) The equilibrium hole concentration \( p_0 = 10^{16} \text{ cm}^{-3} \). Hence, the equilibrium electron concentration

\[ n_0 = \frac{p_0}{e^2} = \frac{1.5 \times 10^{10}}{2.25 \times 10^6} = 6.67 \times 10^3 \text{ cm}^{-3}. \]

Since there is no trapping and \( \tau_n = \tau_p = 5 \ \mu\text{sec} \), therefore, the excess electron and hole concentrations can be given by

\[ \delta n = \delta p = g_{np} \tau_n = g_{np} \tau_p = 5 \times 10^{12} \text{ cm}^{-3}. \]

Therefore, the net electron concentration is given by

\[ n = n_0 + \delta n \approx 5 \times 10^{12} \text{ cm}^{-3}. \]

and the net hole concentration is given by

\[ p = p_0 + \delta p \approx 10^{16} \text{ cm}^{-3}. \]

Therefore, the percentage change in the electron concentration

\[ = \left[ (5 \times 10^{12} - 2.25 \times 10^6) / (2.25 \times 10^6) \right] \times 100\% = 2.22 \times 10^6\%. \]

And the percentage change in the hole concentration \( \approx 0\% \).

Note: with optical excitation (under the low-level injection approximation), there is a very large change in the minority carrier concentration, whereas the change in the majority carrier concentration is hardly noticeable!

(b) Note:

\[ np = 5 \times 10^{12} \times 10^{16} = 5 \times 10^{28} \text{ cm}^{-3} \neq n_1^2. \]

In equilibrium, \( n_0 p_0 = n_1^2 \). Whenever excess carriers are present, \( np \neq n_1^2 \), and the amount of deviation quantifies the departure from equilibrium.

(c) In equilibrium,

\[ E_i - E_F = kT \ln (p_0/n_0) = 0.026 \ln [10^{15} / (1.5 \times 10^{10})] = 0.299 \text{ eV}. \]

In the presence of excess carriers, the electron \( \text{Imref} \)

\[ F_n - E_i = kT \ln (n/n_0) = 0.026 \ln [5 \times 10^{12} / (1.5 \times 10^{10})] = 0.151 \text{ eV} \]

and the hole \( \text{Imref} \)

\[ E_i - F_p = kT \ln (p_0/n_0) = 0.026 \ln [10^{16} / (1.5 \times 10^{10})] = 0.299 \text{ eV}. \]

Thus, the majority carrier \( \text{Imref} \) almost coincides with the equilibrium Fermi level, whereas the minority carrier \( \text{Imref} \) shows a large departure from the equilibrium value.
Photoconductive Devices

- Devices which change their resistance while exposed to light.
- Examples: automatic night light controllers, exposure meters in cameras, moving-object counters, burglar alarms, detectors in fiber optic communication systems, etc.
- Considerations in choosing a photoconductor for a given application: sensitive wavelength range, time response, and optical sensitivity (responsivity) of the material.
- The photoconductivity change while exposed to light is

\[ \Delta \sigma = q\sigma_0 (\tau_n/h + \tau_p/h) \]  

(3.10)

- Obvious that for large changes in \( \Delta \sigma \), the carrier mobility and lifetime should be high (e.g., in \( \ln \text{Sb} \), \( \mu \approx 10^5 \text{ cm}^2/\text{V-sec} \), and could be used as infrared detector with high sensitivity).
- Time response is limited by recombination times, degree of carrier trapping, and time required for the carriers to drift through the device in an electric field.
- Dark resistance (i.e., the resistance of the device without any illumination) should be as small as possible.
- Generally, all these requirements cannot be satisfied simultaneously, and some kind of optimization is required.

Diffusion of Carriers

- When excess carriers are created in a semiconductor and their concentrations vary with position, then there is a net carrier motion from regions of higher concentration to regions of lower concentration.
- This type of motion is called the diffusion, and it is an important charge transport mechanism in semiconductors.
- Diffusion and drift are the two main current transport mechanisms.

Diffusion processes

- Natural result of the random motion of individual electrons.
- Electrons move randomly and experience collisions, on the average, after each mean free time.
Since the motion is truly random, an electron has equal probability of moving into or out of a volume through a boundary.

Fig. 3.5 Spreading of a narrow pulse of electrons created at x = 0 at t = 0 with time ($t_3 > t_2 > t_1$).

A pulse of excess electrons injected at x = 0 at time t = 0 will spread out in time due to diffusion, and eventually $n(x)$ becomes a constant, when no more net motion takes place.

Fig. 3.6 An arbitrary electron distribution along the x-direction: (a) each segment divided into lengths equal to the mean free path $\bar{l}$, and (b) expanded view of two segments centered at $x_0$.

Consider any arbitrary distribution $n(x)$, with x divided into segments $\bar{l}$ (mean free path) wide, with $n(x)$ evaluated at the center of each segment.
• In $\tilde{t}$, half of the electrons in segment (1) to the left of $x_0$ would move into segment (2), and in the same time, half of the electrons in segment (2) to the right of $x_0$ would move into segment (1).

• Therefore, the net number of electrons moving from segment (1) to segment (2) through $x_0$, within a mean free time $\tilde{t} = (n_1 - n_2)\frac{\tilde{I}}{A \sqrt{2}}$, where $A$ is the area perpendicular to $x$.

• Thus, the electron flux density in the $+x$-direction

$$\phi_n(x_0) = \frac{\tilde{I}}{2\tilde{t}} (n_1 - n_2)$$  \hspace{1cm} (3.11)

• Note: is a small differential length, thus,

$$n_1 - n_2 = \frac{n(x) - n(x + \Delta x)}{\Delta x} \bigg|_{\Delta x}$$  \hspace{1cm} (3.12)

where $x$ is taken at the center of segment (1) and $x = \tilde{I}$.

• In the limit of small $\Delta x$ (i.e., small mean free path $\tilde{t}$ between collisions)

$$\phi_n(x) = \frac{\tilde{I}^2}{2\tilde{t}} \lim_{\Delta x \to 0} \frac{n(x) - n(x + \Delta x)}{\Delta x} = -\frac{\tilde{I}^2}{2\tilde{t}} \frac{dn(x)}{dx}$$  \hspace{1cm} (3.13)

• The quantity $\frac{\tilde{I}^2}{2\tilde{t}}$ is called the electron diffusion coefficient $D_n$ ($cm^2/sec$).

• The minus sign in the expression for $\phi_n(x)$ implies that the diffusion proceeds from higher electron concentration to lower electron concentration.

• Similarly, holes diffuse from a region of higher concentration to a region of lower concentration with a diffusion coefficient $D_p$.

Thus,

$$\phi_n(x) = -D_n \frac{dn(x)}{dx} \quad \text{and} \quad \phi_p(x) = -D_p \frac{dp(x)}{dx}$$  \hspace{1cm} (3.14)

and the diffusion current density

$$J_n(\text{diff.}) = +qD_n \frac{dn(x)}{dx} \quad \text{and} \quad J_p(\text{diff.}) = -qD_p \frac{dp(x)}{dx}$$  \hspace{1cm} (3.15)

• Note: electrons and holes move together in a carrier gradient, however, the resulting currents are in opposite directions because of the opposite charges of the particles.

**Diffusion and Drift of Carriers: Built-in Fields**

• The total current density can thus be written as
The total current density is given by:

\[ J_n(x) = q\mu_n n(x)E(x) + qD_n \frac{dn(x)}{dx} \]  
\[ J_p(x) = q\mu_p p(x)E(x) - qD_p \frac{dp(x)}{dx} \]  

and the total current density is

\[ J(x) = J_p(x) + J_n(x) \]  

- The total current may be due primarily to one of these two components, depending upon the carrier concentrations, their gradients, and the electric field.
- Thus, minority carriers can contribute to current conduction significantly through diffusion, even though they contribute very little to the drift term (due to their low concentrations).
- Since electrons drift opposite to the direction of the electric field (due to their negative charge), their potential energy increases in the direction of the electric field.
- The electrostatic potential varies in the opposite direction, and can be given by \[ V(x) = \frac{E_i(x)}{-q} \].

Thus, the electric field can be given by

\[ E(x) = -\frac{dV(x)}{dx} = \frac{1}{q} \frac{dE_i(x)}{dx} \]  

- Note: electrons drift downhill in a band diagram, therefore, the electric field points uphill in a band diagram.
- Note: in equilibrium, no net current => any fluctuation in carrier concentration which brings about a diffusion current also sets up an electric field, which opposes the diffusive motion => thus, equilibrium is established.
- This field is referred to as the built-in field, and can be caused by doping gradients and/or variation in the band gap.
- Equating the hole current density equation to zero, we get

\[ E(x) = \frac{D_p}{\mu_p} \frac{1}{p(x)} \frac{dp(x)}{dx} = \frac{D_p}{\mu_p} \frac{1}{kT} \left( \frac{dE_i}{dx} - \frac{dE_F}{dx} \right) \]  

- Now, \( EF \) does not vary with \( x \) in equilibrium, and the variation of \( E_i \) with \( x \) is given above, thus,

\[ \frac{D}{\mu} = \frac{kT}{q} \]  

(3.21)
• This is an extremely important equation valid for both carrier types, and is called the Einstein relation gives the relation between $D$ and $\phi$, which is a function only of temperature, and allows calculation of one if the other is known.

EXAMPLE 3.3

An intrinsic Si sample is doped with acceptors from one side such that $N_a(x) = N_0(1 + x/L)$, where $N_0 = 10^{15}$ cm$^{-3}$ and $L = 5 \mu$m

(a) Find an expression for $E(x)$ at equilibrium from $x = 0$ to 5 m.
(b) Evaluate $E(x)$ at $x = 0$ and 5 m.
(c) Sketch a band diagram and indicate the direction of $E$

SOLUTION

(a) Recall, at equilibrium, the hole current density $J_p = q\mu_p(x)E(x) - qD_p(d\rho(x)/dx) = 0$ Thus, $E(x) = (D_p/\mu_p)(d\rho(x)/dx)/\rho(x)] = [kT/(q(x + L))]$, where use has been made of the Einstein relation and 100% ionization is assumed. Thus, the electric field varies inversely with distance and has positive values throughout.

(b) $E(0) = 52$ V/cm and $E(L) = 26$ V/cm (c) Note: $E(x) = (1/q)(dE_i/dx)$. Since $E(x)$ varies inversely with $x$, hence $E_i$ (and consequently, both $E_C$ and $E_V$) will have a logarithmic (ln) dependence on $x$.

Diffusion and Recombination: The Continuity Equation

• In the description of conduction processes, the effects of recombination must be included, since they can cause a variation in the carrier distribution.
• Hole conservation equation:

$$\frac{\partial \rho}{\partial t} \bigg|_{x \rightarrow x+\Delta x} = \frac{1}{q} \frac{J_p(x) - J_p(x + \Delta x)}{\Delta x} - \frac{\delta \rho}{\tau_p}$$

(3.22)

i.e., rate of hole buildup = increase of hole concentration in $\Delta x A$ per unit time recombination rate.
Fig. 3.7 Setup to obtain particle count: current entering and leaving a volume $\Delta x A$.

- For $\Delta x \to 0$,

$$\frac{\partial \delta p(x,t)}{\partial t} = \frac{\partial \delta p}{\partial t} = -\frac{1}{q} \frac{\partial J_p}{\partial x} - \frac{\delta p}{\tau_p} \quad (3.23)$$

- This is called the continuity equation for holes, and, similarly, for electrons, we can write

$$\frac{\partial \delta n}{\partial t} = \frac{1}{q} \frac{\partial J_n}{\partial x} - \frac{\delta n}{\tau_n} \quad (3.24)$$

- When the current is carried entirely due to diffusion (negligible drift), then we obtain the diffusion equation for electrons, given by

$$\frac{\partial \delta n}{\partial t} = D_n \frac{\partial^2 \delta n}{\partial x^2} - \frac{\delta n}{\tau_n} \quad (3.25)$$

and, similarly for holes,

$$\frac{\partial \delta p}{\partial t} = D_p \frac{\partial^2 \delta p}{\partial x^2} - \frac{\delta p}{\tau_p} \quad (3.26)$$

**Steady State Carrier Injection: Diffusion Length**

- In steady state, if a distribution of excess carriers is maintained, then the diffusion equations become

$$\frac{d^2 \delta n}{dx^2} = \frac{\delta n}{L_n^2} \quad \text{and} \quad \frac{d^2 \delta p}{dx^2} = \frac{\delta p}{L_p^2} \quad (3.27)$$

where $L_n (= \sqrt{D_n \tau_n})$ and $L_p (= \sqrt{D_p \tau_p})$.
are the electron diffusion length and the hole diffusion length, which is the average distance an electron or a hole diffuses before recombining respectively.

- **Case study:** suppose excess holes are injected into a semi-infinite n-type bar, which maintains a constant concentration \( \delta p(x = 0) = \Delta p \) (relevant problem in a forward biased diode).

- Obviously, the excess holes would diffuse into the n-type bar, recombine with the electrons with a characteristic lifetime \( \tau_p \) (and diffusion length \( L_p \)), and for large values of \( x \), the excess hole concentration should decay to zero; thus,

\[
\delta p(x) = \Delta p e^{-x/L_p} \tag{3.28}
\]

and the decay profile is exponential.

- The steady state distribution of excess holes causes diffusion, and therefore, a hole current in the direction of decreasing concentration, given by

\[
J_p(x) = -qD_p \frac{dp}{dx} = -qD_p \frac{d\delta p}{dx} = q \frac{D_p}{L_p} \Delta p e^{-x/L_p} = \frac{qD_p}{L_p} \delta p(x) \tag{3.29}
\]

(This equation would come handy in the diode analysis.)

**The Haynes-Shockley Experiment**

- Counterpart of the Hall effect experiment.
- Independently determines the minority carrier mobility and diffusion coefficient

![Fig.3.8 Schematic for Haynes-Shockley experiment: drift and diffusion of a hole pulse in an n-type bar.](mywbut.com)
(a) sample geometry,  
(b) shape and position of the pulse for different times along the bar.

**Basic principle**

- a pulse of excess holes is created in an n-type bar (which has an applied electric field),
- as time progresses, the holes spread out by diffusion and move due to the electric field, and their motion is monitored somewhere down the bar,
- the time required for the holes to move the distance gives a measure of their mobility, and the spreading of the pulse in a given time gives a measure of the diffusion coefficient.
- A pulse of excess carriers is created by a light flash at $x = 0$ in an n-type semiconductor bar with an electric field $E$.
- The excess holes drift down the bar and reach the point $x = L$ after a time $t_d$, thus, the drift velocity $v_d = L/t_d$, and the hole mobility $\mu_p = v_d/E$.
- For measurement of diffusion coefficient, assume the pulse spreads without drift and neglect recombination; then the diffusion equation can be rewritten as

\[
\frac{\partial \delta p(x,t)}{\partial t} = D_p \frac{\partial^2 \delta p(x,t)}{\partial x^2} 
\]  

(3.30)

- The solution to this equation is the well known Gaussian distribution, given by

\[
\delta p(x,t) = \left[ \frac{\Delta P}{2\sqrt{\pi D_p t}} \right] e^{-x^2/4D_pt}
\]  

(3.31)

where the number of holes per unit area created over a negligibly small distance at $t = 0$.

- Note: the peak values of the pulse decreases and the pulse spreads in $\pm x$ directions with time.
- Let the peak value of the pulse be at $\delta p$ at $t = t_d$ and note that at $x = \Delta x/2$, $\delta p$ is down by $1/e$ of its peak value $\delta p$; thus,

\[
e^{-1}\delta p = \delta p e^{-\left(\Delta x/2\right)^2/4D_pt_d}
\]  

(3.32)

and $D_p = (\Delta x^2)/16t_d$, where $\Delta x = \Delta t v_d = \Delta t L/t_d$, where $t$ is the spread of the pulse seen in an oscilloscope in time.
**EXAMPLE** A p-type Si sample is used in the Haynes-Shockley experiment. The length of the sample is 2 cm and the two measurement probes are separated by 1.9 cm. The voltage applied across the two ends of the sample is 5 V. A pulse arrives at the collection point 0.608 msec after its injection at the injection point, and the spread of the pulse t at the collection point is 180 sec. Calculate the electron mobility and diffusion coefficient, and verify whether Einstein relation is satisfied.

**SOLUTION**

The electron mobility

\[
\mu_n = \frac{v_a}{E} = \frac{1.9/(0.608 \times 10^{-3})}{5/2} = 1250 \text{ cm}^2/\text{V-sec}
\]

The electron diffusion coefficient

\[
D_n = \frac{(\Delta x)^2}{16t_d} = \frac{(\Delta t)^2}{16t_d^3} = \frac{(180 \times 10^{-6} \times 1.9)^2}{16 \times (0.608 \times 10^{-3})^3} = 32.5 \text{ cm}^2/\text{sec}
\]

Their ratio

\[
\frac{D_n}{\mu_n} = \frac{32.5}{1250} = 0.026 = \frac{kT}{q}
\]

Thus, Einstein relation is indeed satisfied.
• Any combination of drift and diffusion implies a gradient in the steady state Imrefs.
• Under general case of nonequilibrium electron concentration with drift and diffusion, the total electron current can be written as

\[ J_n(x) = q\mu_n n(x)E(x) + qD_n \frac{dn(x)}{dx} \]  \hspace{1cm} (3.33)

• Using the expression for \( n(x) \) in terms of the electron Imref, and applying Einstein relation, it can be shown that

\[ J_n(x) = \mu_n n(x) \frac{dF_n}{dx} \]  \hspace{1cm} (3.34)

and, similarly, for holes,

\[ J_p(x) = \mu_p p(x) \frac{dF_p}{dx} \]  \hspace{1cm} (3.35)

• Therefore, any drift, diffusion, or a combination of the two in a semiconductor sets up currents proportional to the gradient in the Imrefs, or, in other words, no current implies constant Imrefs.

**Practice Problems**

3.1 A 100 mW laser beam with wavelength \( \lambda = 632.8 \text{ nm} \) is focused onto an InP (\( E_g = 1.35 \text{ eV} \)) sample 100 \( \mu \text{m} \) thick. The absorption coefficient at this wavelength is \( 10^6 \text{ cm}^{-1} \). Find the number of photons emitted per second by radiative recombination in the sample, assuming 100% quantum efficiency (i.e., each incident photon creates one EHP, and they spontaneously recombine). What power is delivered to the sample as heat?

3.2 A photon of monochromatic light of wavelength 500 nm is absorbed in GaAs (\( E_g = 1.42 \text{ eV} \), \( m_n^* = 0.067 m_0 \), and \( m_p^* = 0.48 m_0 \)) and excites an electron from the valence band into the conduction band. Calculate the kinetic energies of the electron and the hole.

3.3 Starting from the recombination/generation rate equation, determine the excess electrons [created in a p-type sample (with equilibrium carrier concentrations given by \( p_0 \) and \( n_0 \)) by a high intensity pulse of light] decay profile as a function of time [i.e., \( \delta n(t) \)]. Assume high-level injection condition (i.e., \( \delta n(t) \gg p_0, n_0 \)), and at \( t = 0 \), \( \delta n(0) = \Delta n \). State and justify whether \( \delta n(t) \) would decay with the same profile till it reaches zero.

3.4 A Ge sample (\( n_i = 2.33 \times 10^{13} \text{ cm}^{-3} \), \( \mu_n = 3900 \text{ cm}^2/\text{V-sec} \), and \( \mu_p = 1900 \text{ cm}^2/\text{V-sec} \)) with \( n_0 = 10^{17} \text{ cm}^{-3} \) is optically excited at 300 K such that \( g_{\text{opt}} = 10^{20} \text{ EHP/cm}^2/\text{sec} \), and \( \tau_n = \tau_p = 10 \). What is the separation of the Imrefs \( (F_n - F_p) \)? Clearly draw the band diagram showing the Imrefs and the equilibrium Fermi level \( E_F \). Also, compute the change in the sample conductivity after illumination.
3.5 A sample of p-type Si has a dark resistivity of $10^{12} \Omega \cdot \text{cm}$ at 300 K. The sample is illuminated uniformly to generate $10^{19} \text{ EHP/cm}^2 \cdot \text{sec}$. The electron lifetime in the sample is 1 $\mu\text{sec}$. Calculate the sample resistivity and the percent change in the conductivity after illumination due to the majority and the minority carriers.

3.6 Light is shone uniformly on a 0.5 $\Omega \cdot \text{cm}$ n-type Si sample for a long time to attain steady-state, and the difference between the electron and the hole Imrefs is found to be 0.55 eV. Now, the light is suddenly shut off at some arbitrary time (call that $t = 0$), and the excess conductivity is found to decrease to 10% of its maximum value at time $t = 40 \mu\text{sec}$. Determine the optical generation rate $g_{op}$ and the excess hole lifetime $\tau_p$. Assume low-level injection and no trapping.

3.7 A sample is doped with donors such that $n_0 = \frac{G}{L} \exp(-x/L)$, where $G$ is a constant, $L$ is the length of the sample, and $n_0$ is always $\gg n_i$. Assuming equilibrium, find the built-in electric field in order to sustain this distribution, and clearly draw the band diagram. Also, plot the potential $V(x)$ as a function of position.

3.8 A 4.63 $\Omega \cdot \text{cm}$ type Si sample is illuminated uniformly at $t = 0$ to produce EHPs. Starting from the continuity equation and assuming low-level injection and no current flow, determine the expression for the build-up of excess holes as a function of time. If the excess conductivity at $t = 1 \mu\text{sec}$ is $0.293 (\text{k} \Omega \cdot \text{cm})^{-1}$, and after sufficiently long time, it is $2.93 (\text{k} \Omega \cdot \text{cm})^{-1}$, determine the optical generation rate $g_{op}$ and the excess hole lifetime $\tau_p$. Assume no trapping.

3.9 The following data are obtained from the Haynes-Shockley experiment on a p-type Si sample at 300 K: length of sample=2cm, length between injection and collection probes =1.2 cm, applied voltage=10 V, $t_i = 150 \mu\text{sec}$, and $\Delta t = 39.5 \mu\text{sec}$. Calculate the mobility and diffusion coefficient of the minority carriers, and check if this data satisfies the Einstein relation. What should be the minimum values of the lifetime and the diffusion length in the original sample for authentic measurement results?

3.10 In the Haynes-Shockley experiment discussed in this chapter, the recombination of the excess carriers was neglected. However, by a simple modification, it can be made to include the effects of recombination. Assume an n-type semiconductor, the peak voltage of the pulse displayed on the CRO screen is proportional to the peak value of the hole concentration under the collector terminal at time $t_d$, and that the displayed pulse can be approximated as a Gaussian, which decays due to recombination by $\exp(-t/\tau_p)$, where $\tau_p$ is the excess hole lifetime. The electric field is varied and the following data taken: for $t_d = 200 \mu\text{sec}$, the peak is 20 mV; and for $t_d = 50 \mu\text{sec}$, the peak is 80 mV. What is $\tau_p$?