

Physics of Semiconductor Devices in Equilibrium

Review of Quantum Mechanics

On the atomic scale, deterministic classical mechanics breaks down and quantum mechanics is required. However, we will be able to use quantum mechanics to derive approximations that can be used in treating electrons in semiconductors as quasi-classical particles under many conditions.

Begin with the **time-dependent** Schrödinger Equation, which is the fundamental relation describing matter on the quantum scale:

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\Psi + V(\vec{x})\Psi = -\frac{\hbar}{j}\frac{\partial\Psi}{\partial t} \quad (1)$$

$$\hbar = \frac{h}{2\pi}$$

$V(\vec{x})$: potential field for electron

$\Psi(\vec{x},t)$: state function, solution to S's equation

Assume we can separate variables:

$$\Psi(\vec{x},t) = \psi(\vec{x})\phi(t)$$

Substituting into (1) and gathering time and position terms together:

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\vec{\nabla}^2\psi + V = -\frac{\hbar}{j}\frac{1}{\phi}\frac{\partial\phi}{\partial t} \quad (2)$$

Since the left side depends only on position and the right only on time, for them to always be equal they must both equal a constant. That constant is E , the total energy.

The time dependence is simple to calculate:

$$\frac{\partial\phi}{\partial t} = -\frac{jE}{\hbar}\phi \quad (3)$$

so

$$\phi(t) = Ae^{-(2\pi jE/h)t} = Ae^{-(2\pi j\nu)t} \quad (4)$$

since $E = h\nu$.

The problem then becomes solving the **time-independent** Schrödinger Equation,

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\psi + V\psi = E\psi \quad (5)$$

The solutions depend on the form of the potential V and the boundary conditions. In general, only certain values of E (eigenvalues) allow solutions to the equation. Associated with the allowed energy levels are wavefunctions ψ (eigenfunctions). Together they describe the possible states of the system.

We can also compare to the classical case:

$$E = KE + V = \frac{|\vec{p}|^2}{2m} + V \quad (6)$$

We can make a direct analogy by using operators for the momentum:

$$\vec{p} \rightarrow \frac{\hbar}{j}\vec{\nabla} \quad (7)$$

and the energy:

$$E \rightarrow -\frac{\hbar}{j}\frac{\partial}{\partial t} \quad (8)$$

In contrast to a classical system, the solution to Schrödinger's equation only gives a probability density for the electron location.

$$|\Psi(\vec{x}, t)|^2 : \text{probability of finding electron at position } \vec{x} \text{ at time } t$$

We can describe a normalization condition, since the electron exists somewhere:

$$\int_{-\infty}^{\infty} \psi^* \psi d\vec{x} = 1$$

ψ and its gradient must also be continuous since it describes a physical system.

Because of the probabilistic nature of ψ , only expected values for parameters such as position or momentum can be determined. For example,

$$\langle \vec{x} \rangle = \int_{-\infty}^{\infty} \psi^* \vec{x} \psi d\vec{x}$$

$$\langle \vec{p} \rangle = \int_{-\infty}^{\infty} \psi^* \frac{\hbar}{j} \vec{\nabla} \psi d\vec{x}$$

Atomic Structure

An example of a potential for which the allowed energy levels and associated wavefunctions can be determined is for atomic potentials. For the hydrogen atom (which can be done exactly, see any modern physics text), the allowed electron energies are given by

$$E = \frac{-m_0 q^4}{8h^2 \epsilon_0^2} \frac{1}{n^2} \quad n = 1, 2, 3, \dots \quad (9)$$

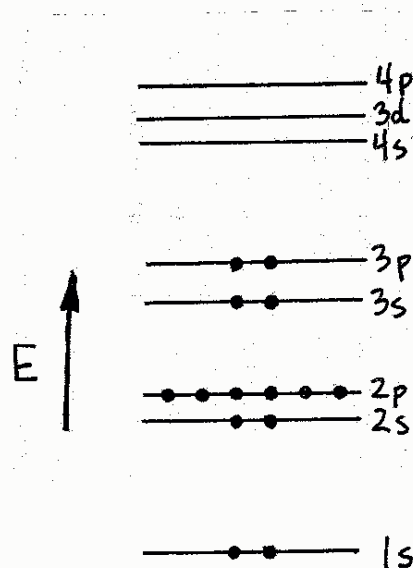
The allowed states and associated wavefunctions can be described by 4 quantum numbers:

1. Principle quantum number: $n = 1, 2, 3, \dots$
2. Angular momentum quantum number: $l = 0, 1, 2, \dots, n - 1$
3. Magnetic quantum number: $m = 0, \pm 1, \pm 2, \dots, \pm l$
4. Electron spin: $s = \pm 1/2$

For hydrogen, the energy depends only on the principal quantum number n , but for multi-electron atoms, where electron-electron interactions appear, the energy levels depend also on l and m , breaking all but the spin degeneracy.

Quantum mechanics further requires that each electron must have a distinct energy state (unique solution) defined by a unique set of quantum numbers (Pauli exclusion principle).

Silicon has its two inner shells (1s, 2s and 2p subshells) totally filled. It's outer shell has 4 electrons, 2 in 3s orbitals and 2 in the 6 3p orbitals.



Energy bands in solids

Materials may be classified as:

1. Amorphous: no long-range structure (i.e., glass)
2. Polycrystalline: many small ordered regions
3. Crystalline: long-range 3D order, a repeating unit cell

Important semiconductor materials are primarily crystalline. Many of the important properties of semiconductors are a result of the regular crystal lattice structure. The crystal structure determines the potential V seen by electrons in the structure.

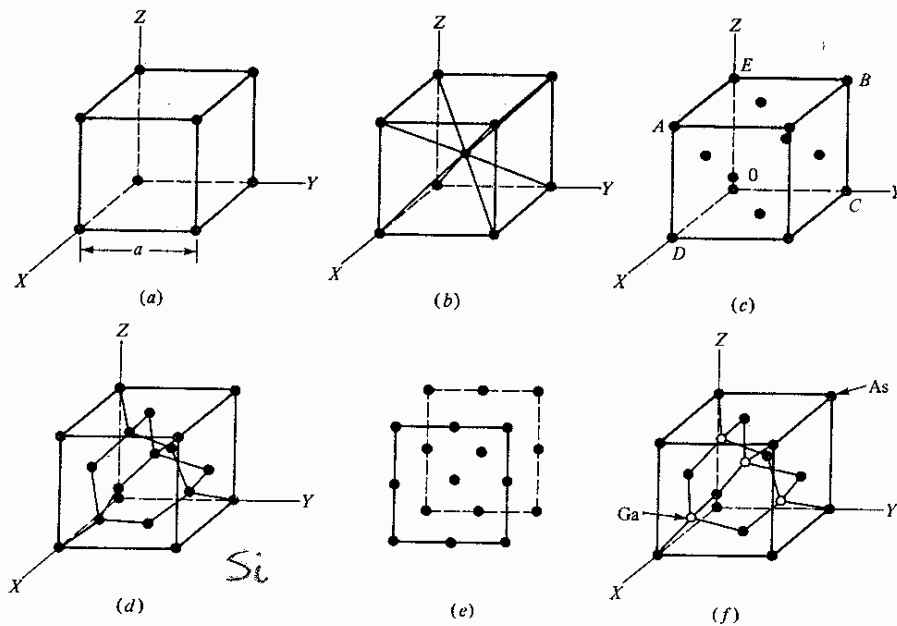


FIGURE 1-3
Unit cells of cubic crystals: (a) simple cubic, (b) body-centered cubic, (c) face-centered cubic, (d) diamond, (e) two penetrating fcc lattices in two dimensions, (f) zinc blende.

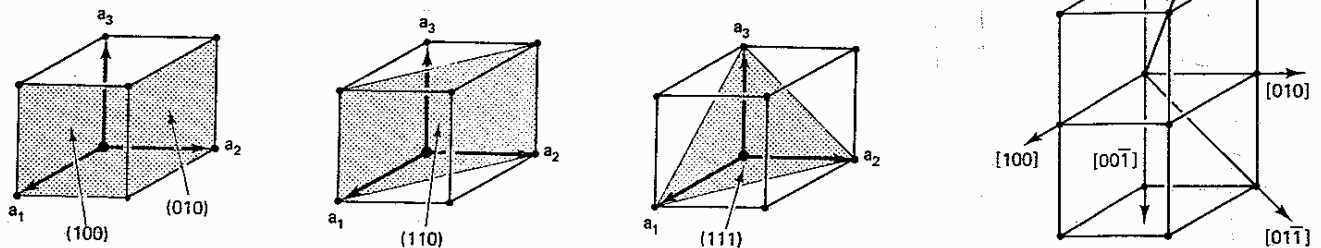
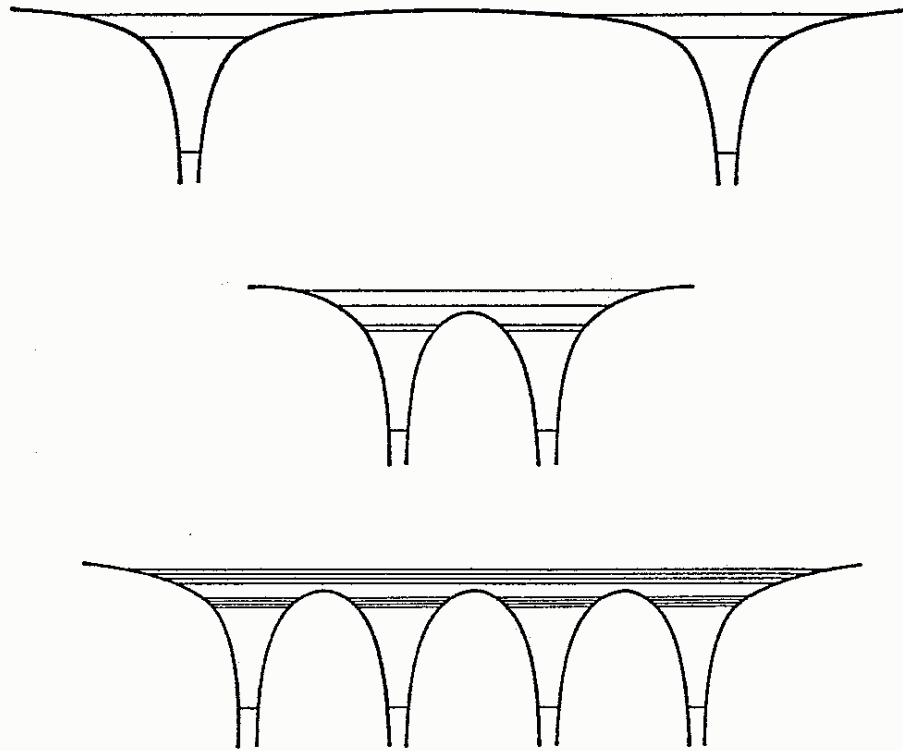


Figure 1.19 Examples of Miller indices for various planes and directions in a cubic lattice.

The fixed position of atoms within the lattice and the bonding between neighbors determines the material properties. The outermost or valence electrons play an essential role in determining those properties.

Effect of Lattice.

Lower energy core states remain localized. Outer electron wavefunctions overlap and are influenced by the potential of neighbors. The outer valence electrons become delocalized and by Pauli exclusion principle, the levels split into N sublevels (N is number of atoms in crystal). For a large crystal with N large, an essentially continuous band of allowed energies is formed (the valence band).



The strength with which the outer or valence electrons are bound to the atoms determines the properties of the material. There are three basic types of materials distinguished by large differences in electronic conduction.

Conductor: Valence electrons are not bound but are free to move throughout the solid. Excellent conductor: $\rho \sim 10^{-5} \Omega\text{-cm}$ in metals.

Insulator: Valence electrons are tightly bound to the nuclei, not available for conduction: $\rho \sim 10^{16} \Omega\text{-cm}$ in SiO_2 .

Semiconductor: Valence electrons shared in covalent bonds, less tightly bound, some may become free due to thermal energy: $\rho \sim 10^5 \Omega\text{-cm}$ in pure Si at room temperature (much smaller in doped material).

Example of a Covalent Semiconductor: Silicon

By sharing each of its outer electrons with its 4 nearest neighbors, a group IV semiconductor such as silicon can complete its outer shell. These electrons are covalently bonded. They are not tightly bound and (especially at high temperatures) can become free to contribute to electronic conduction. In order to satisfy the Pauli exclusion principle, before electron sharing is possible, one of the 3s electron must move up to a 3p level. This costs some energy that is more than compensated by the reduction due to the electron sharing. The resulting orbitals are described as sp^3 hybrid orbitals and each arises from linear combinations of the s and p orbitals.

One way to visualize what happens as atoms are brought together to form a crystal is to examine theoretical calculations for the allowed energy bands as the silicon lattice distance is reduced, starting with the essentially isolated atoms.

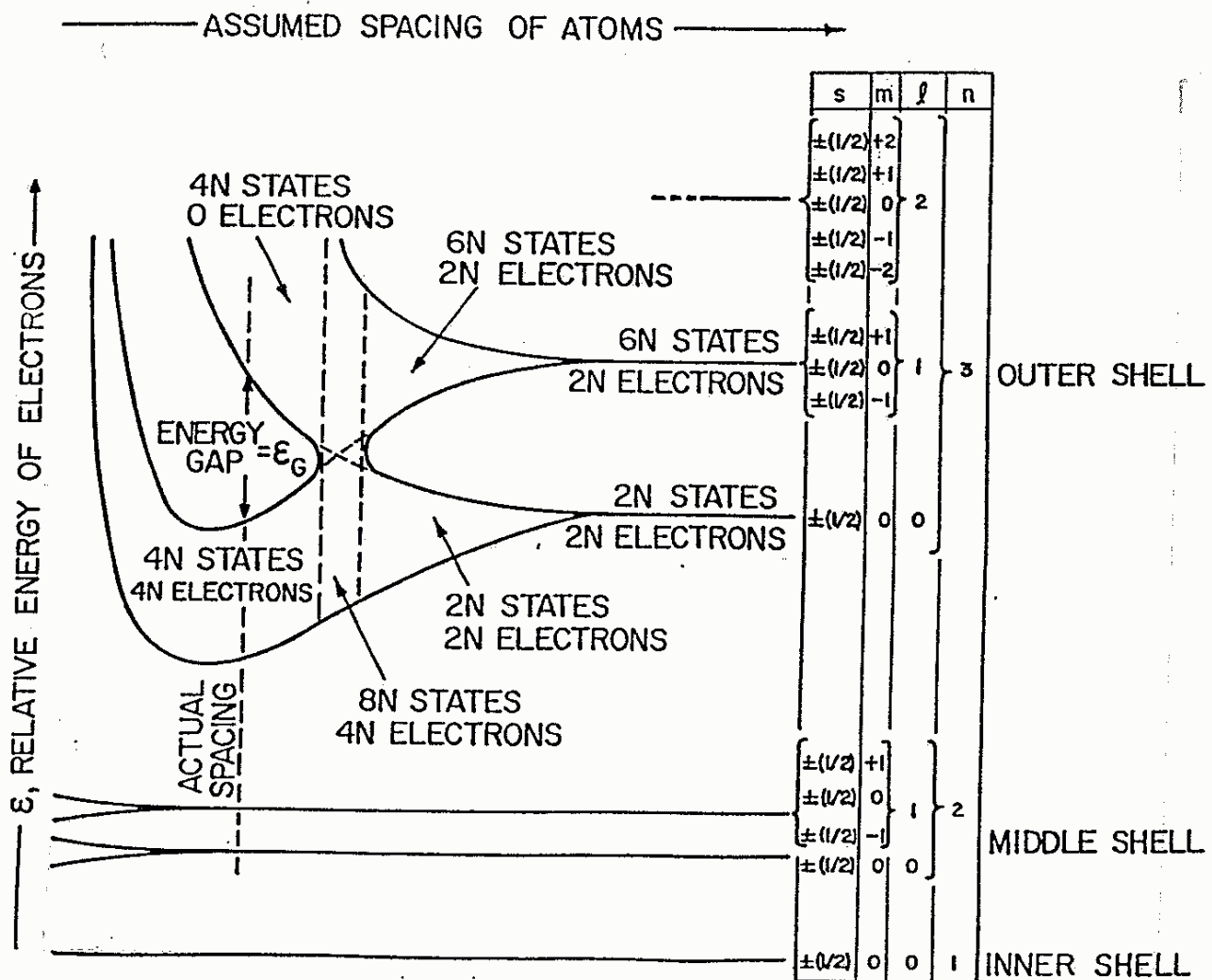


Figure 2-14 Theoretical energy-level scheme vs. atomic spacing for silicon.

Thermal energy breaks bonds liberating electrons and creating a "hole" or vacant state. Both the electron and the hole are mobile and contribute to conduction. Note that in a pure semiconductor $n = p$. Since electrons are only allowed to exist in distinct energy levels, there are allowed bands of energy states. Valence electrons exist in the valence band which is totally full at 0°K since, by sharing, each atom has its outer shell full. The conduction electrons exist in the conduction band which has higher energies.

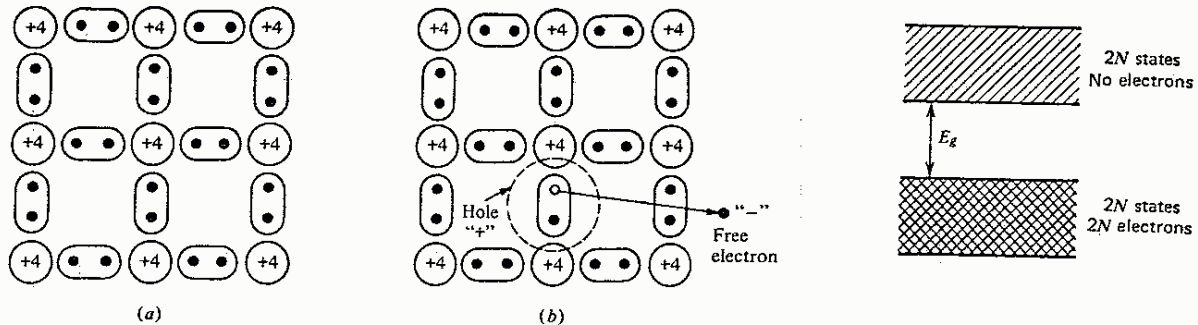


FIGURE 1-8
Two-dimensional schematic representations of crystal structure in silicon: (a) complete covalent bond and (b) broken covalent bond.

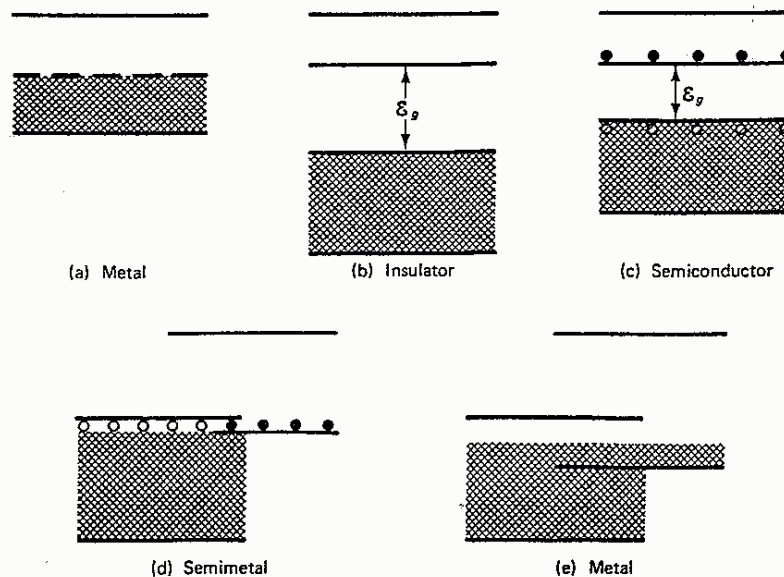
The probability of an electron getting sufficient thermal energy to get into conduction band is:

$$p \simeq 1 \text{ for metals}$$

$$p \simeq 0 \text{ for insulators}$$

$$0 \leq p \leq 1 \text{ for semiconductors}$$

Whether a material is an insulator, conductor or semiconductor can be inferred directly from the band structure.



Example: Free Electron

Let's look at a simple example, a free electron in 1D. For $E > V_0$, we can let

$$k^2 = (2m/\hbar^2)(E - V_0). \quad (10)$$

$$\psi(x) = A \exp(jkx) + B \exp(-jkx) \quad (11)$$

When the time dependence is included, this solution is in the form of traveling waves in the positive and negative direction, with constant amplitude everywhere (the electron is equally likely to be anywhere). $k = 2\pi/\lambda$ is the wavenumber (the wavevector in 3D, $\vec{k} = (k_x, k_y, k_z)$).

$$E - V_0 = \text{Kinetic Energy} = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m} \quad (12)$$

$\hbar \vec{k}$ is called the **crystal momentum**. It will act like momentum in determining the response to external fields and is conserved in phonon or photon interactions, but it is not the true momentum which also responds to internal fields from the periodic potential.

For the free electron, there are no internal crystal fields so $p = \hbar k$:

$$\langle p \rangle = \frac{\int_{-\infty}^{\infty} (A \exp(-jkx)) \frac{\hbar}{j} \frac{\partial}{\partial x} (A \exp(jkx)) dx}{\int_{-\infty}^{\infty} (A \exp(-jkx)) (A \exp(jkx)) dx} = \hbar k \quad (13)$$

Real particle: wavepacket, infinite sum of waves which interfere destructively everywhere but in a localized region. In the form of Fourier integrals:

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k) \exp(jkx) dk$$

$$A(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) \exp(-jkx) dx$$

In order to reduce Δx , the range in k -values (Δk) must be increased. This is a restatement of Heisenberg's uncertainty principle, $\Delta x \Delta k > 1$ and follows directly from the wave nature of matter.

Velocity of particle = group velocity of wave packet envelope. Since $v_g = p/m$ and $KE = (E - V_0) = p^2/2m$:

$$v_g = \frac{\partial(KE)}{\partial p} = \frac{\partial(E - V_0)}{\partial p} = \frac{\partial E}{\partial p} = \frac{1}{\hbar} \frac{\partial E}{\partial k} \quad (14)$$

In 3D:

$$\vec{v}_g = \vec{\nabla}_p E = \frac{1}{\hbar} \vec{\nabla}_k E = \frac{1}{\hbar} \left(\frac{\partial E}{\partial k_x} \vec{x} + \frac{\partial E}{\partial k_y} \vec{y} + \frac{\partial E}{\partial k_z} \vec{z} \right) \quad (15)$$

Equations (14) and (15) are valid in general, not just for a constant potential.

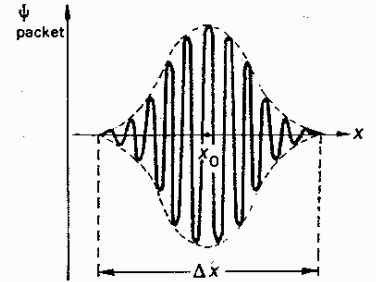


Figure 6.2. A wave packet:
 — Re $\psi(x)$
 ---- $|\psi(x)|$

Nearly Free Electron Approximation:

For a free electron:

$$E = \frac{\hbar^2 k^2}{2m}$$

Bragg reflection: when $n\lambda = 2a$, $n = \pm 1, \pm 2, \dots$, waves are reflected off the crystal lattice planes. Such that for

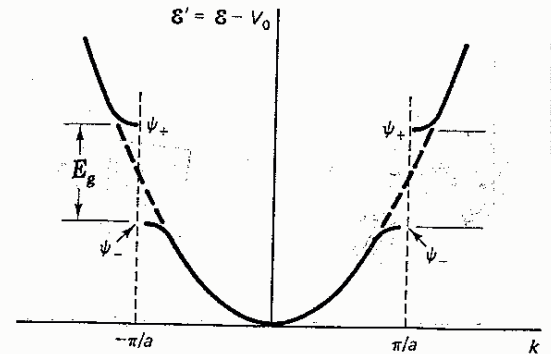
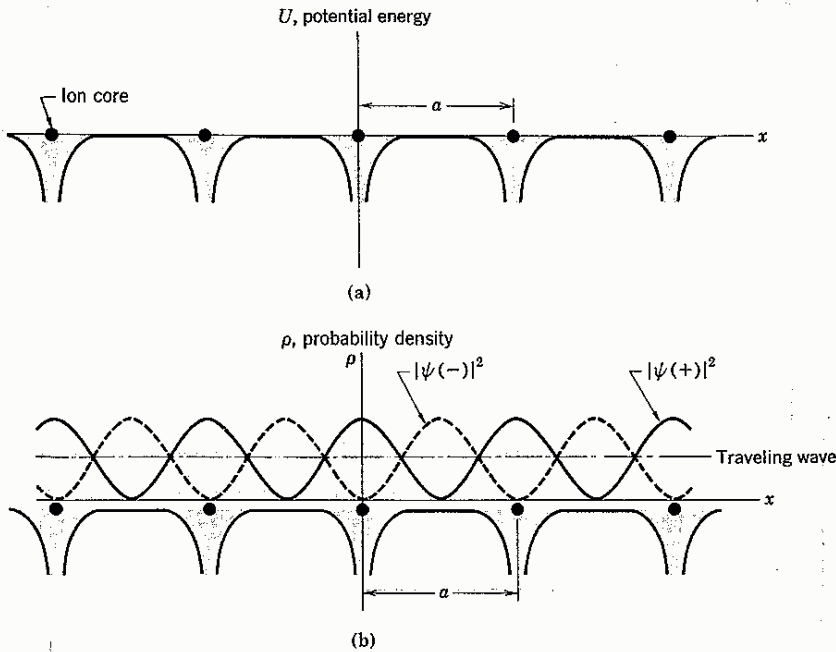
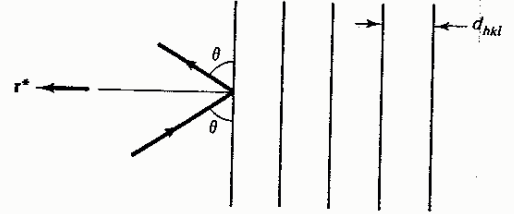
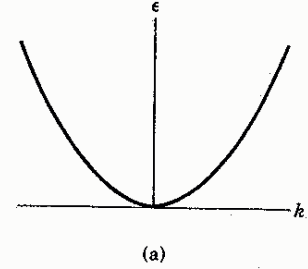
$$k = \frac{2\pi}{\lambda} = \frac{n\pi}{a}$$

the only time-independent states are standing waves rather than travelling waves as in the constant potential case:

$$\psi(+) = A/2 [\exp(jn\pi x/a) + \exp(-jn\pi x/a)] = A \cos(n\pi x/a) \quad (16)$$

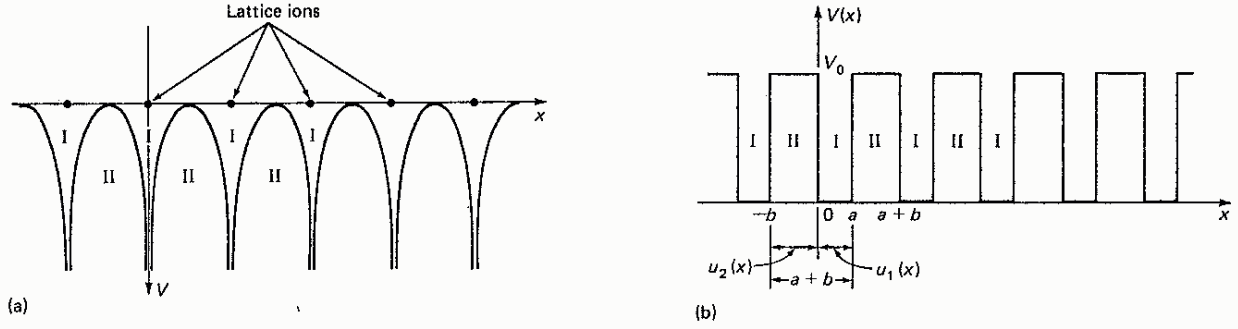
$$\psi(-) = A/2 [\exp(jn\pi x/a) - \exp(-jn\pi x/a)] = Aj \sin(n\pi x/a) \quad (17)$$

While travelling waves have constant probability densities, the two standing waves pile up charge preferentially either near (+) or away from (-) the atom cores. Since the electron potential is less near the positively charged ion cores, the energy of the two states are perturbed in opposite directions from the value they would have in a constant averaged potential. The result is bands of allowed energies divided by energy gaps where no allowed states exist.



Kronig-Penney model

We will look next at a very simple one-dimensional potential. Many of the essential features of energy bands in real three-dimensional crystals appear in the analysis of this simple model.



$$V = V_0; \quad \text{for } -b < x < 0, a < x < a+b, \text{ etc.}$$

$$V = 0; \quad \text{for } 0 < x < a, a+b < x < 2a+b, \text{ etc.}$$

Bloch proved that solutions to Schrödinger's equation in a periodic potential are of the form:

$$\psi(\vec{x}) = u(\vec{x})e^{j\vec{k} \cdot \vec{x}} \quad (18)$$

where $u(x)$ is a function with the periodicity of the lattice.

Assuming a solution of this type in 1D, it is straightforward to find solutions for $u(x)$ in the two types of regions (I and II). Substituting into S's equation using the potential in each of the regions:

$$\frac{\partial^2 u_1}{\partial x^2} + 2jk \frac{\partial u_1}{\partial x} - (k^2 - \alpha^2)u_1 = 0 \quad (\text{In region I}) \quad (19)$$

$$\frac{\partial^2 u_2}{\partial x^2} + 2jk \frac{\partial u_2}{\partial x} - (k^2 - \beta^2)u_2 = 0 \quad (\text{In region II}) \quad (20)$$

where

$$\alpha^2 = \frac{2mE}{\hbar^2} \quad \text{and} \quad \beta^2 = \frac{2m(E - V_0)}{\hbar^2}$$

The solutions in the two regions are:

$$u_1(x) = Ae^{j(\alpha-k)x} + Be^{-j(\alpha+k)x} \quad (\text{In region I}) \quad (21)$$

$$u_2(x) = Ce^{j(\beta-k)x} + De^{-j(\beta+k)x} \quad (\text{In region II}) \quad (22)$$

Each of these solutions has two unknown parameters. These parameters can be determined using the boundary conditions. Since $\psi(x)$ and its derivative must be continuous, the same may be said for $u(x)$. Applying these conditions at the two interfaces between the regions (i.e., at 0 and a) gives four linear equations. The normalization condition must also be satisfied. Solutions exist only if the four boundary condition equations are linearly dependent. Using linear algebra, it can be shown that a solution exists only if

$$-\left[(\alpha^2 + \beta^2)/2\alpha\beta\right] \sin \alpha a \sin \beta b + \cos \alpha a \cos \beta b = \cos [k(a+b)]. \quad (23)$$

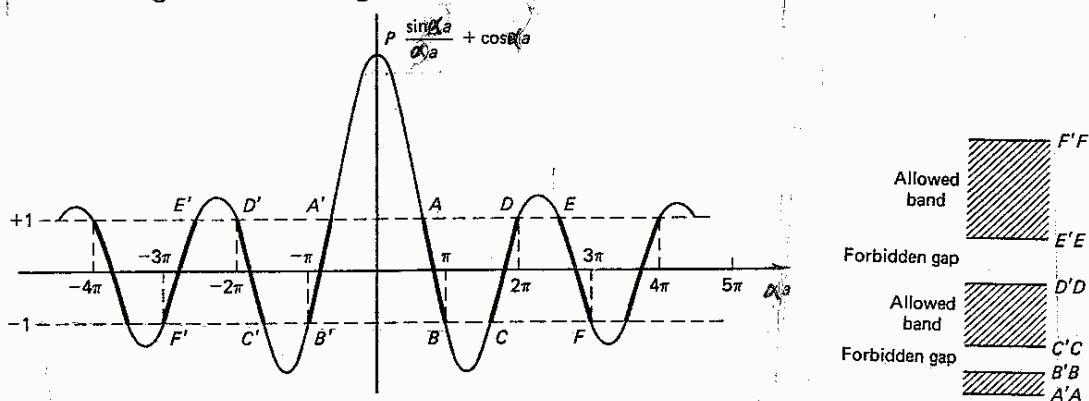
α and β are known functions of E , so this represents a transcendental relationship between E and k which could be solved (graphically or numerically).

It turns that out since the right hand side is always between ± 1 , only certain values of E allow solutions. We can see this more clearly if we do a further simplification. Assume that $V_0 \rightarrow \infty$ while $b \rightarrow 0$, keeping the product $V_0 b$ constant. In that case (23) reduces to

$$P \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos ka \quad (24)$$

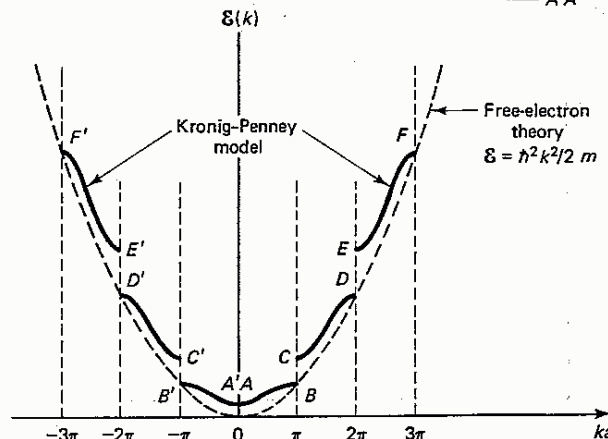
where $P = mV_0ba/\hbar^2$.

Once again, the right side is limited to values in the range ± 1 , so only certain values of α and thus E allow solutions. This gives alternating bands of allowed and disallowed energy values.



An energy versus wavevector (E vs k) relationship as shown to the right can also be calculated.

As the potentials get stronger (P increases) the bands get narrower, approaching discrete levels like in an isolated atom. As the potential variation gets weaker, the bands broaden until eventually all energies are allowed as in the case of the free electron.



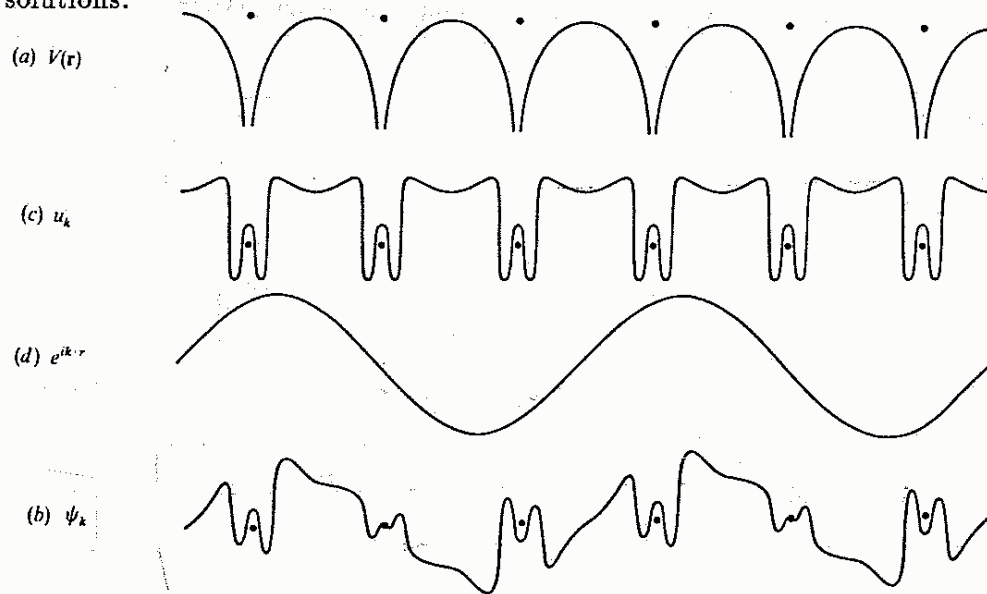
Extended and reduced zone scheme

As stated previously, in a periodic potential, the solutions to Schrödinger's equation are Bloch functions:

$$\psi(\vec{x}) = u(\vec{x})e^{j\vec{k}\cdot\vec{x}} \quad (25)$$

where $u(x)$ is a function with the periodicity of the lattice.

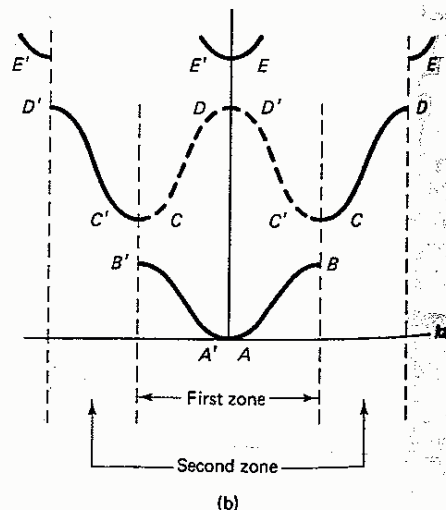
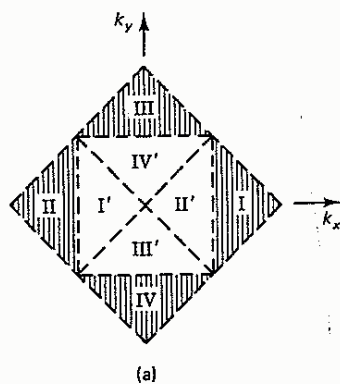
Below is an example of how the periodic function and the plane wave combine to give the wavefunction solutions:



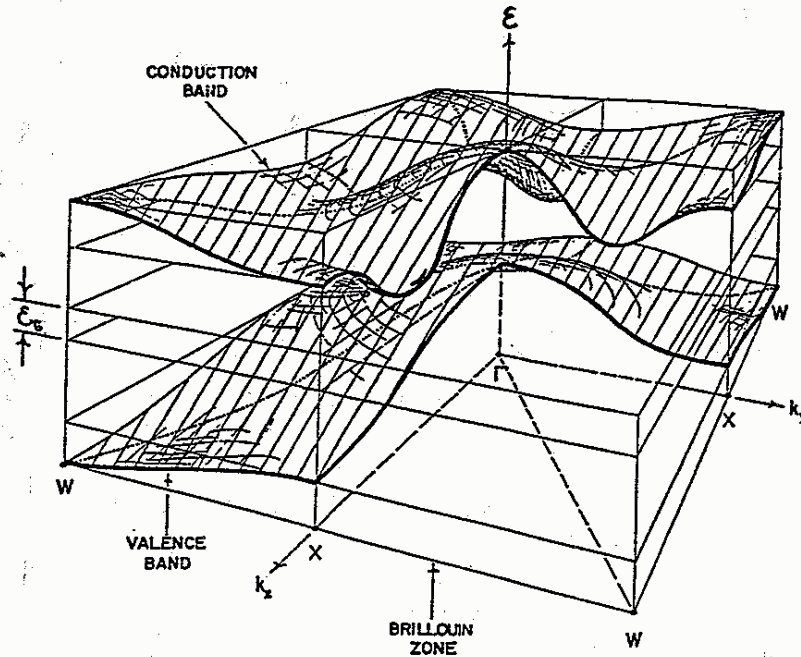
One of the properties of Bloch function solutions is their periodicity in \vec{k} . In 1D, let $k' = k + 2n\pi/a$ and $\psi_{k'}(x) = u_{k'}(x) \exp jk'x$

$$\psi_{k'}(x) = u_{k'}(x) \exp jk'x = [u_{k'}(x) \exp (-j2n\pi x/a)] \exp jkx = u_k(x) \exp jkx = \psi_k(x) \quad (26)$$

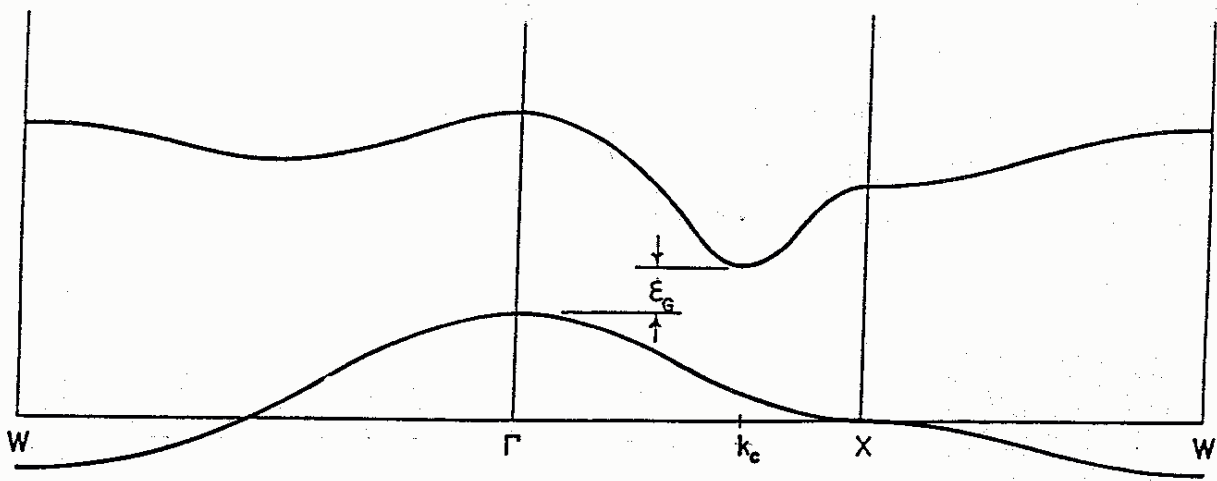
Thus, the wavefunction for k' and k are identical if k' and k differ by $n\pi/a$ (a reciprocal lattice vector in 3D). We can therefore only consider the energy bands for \vec{k} near the origin. This is called the reduced zone scheme. All larger values can be mapped into that region of k -space which is called the first Brillouin zone.



So far we have dealt primarily with one dimension. For a real crystal, the periodicity of the lattice and thus the Bragg reflections change as a function of direction. Therefore, the shape of the energy bands will also change depending on the direction of \vec{k} .



A 3D band structure is often represented in 1D by looking only at the directions of highest symmetry where the maxima and minima generally lie. Those maxima and minima do not in general occur at the same value of \vec{k} .



Response of electrons to external forces.

We saw before that,

$$v_g = \frac{1}{\hbar} \frac{\partial E}{\partial k} \text{ or } v_g = \frac{1}{\hbar} \vec{\nabla}_k E \text{ in 3D} \quad (27)$$

For a constant force, $F = q\mathcal{E}$ and $dE/dt = Fv_g$ (work done by a force).

$$Fv_g = \frac{dE}{dt} = \frac{\partial E}{\partial k} \frac{dk}{dt} = \hbar \frac{dk}{dt} v_g \quad (28)$$

$$F = \hbar \frac{dk}{dt} = \frac{d(\hbar k)}{dt} = \frac{dp}{dt} \quad (29)$$

Compare to classical:

$$F = ma = m \frac{dv}{dt} = \frac{d(mv)}{dt} = \frac{dp}{dt} \quad (30)$$

$\hbar k$ is crystal momentum and includes only the effects of external fields not internal (crystal lattice) forces. The effect of the lattice potential is reflected in the E vs. k diagram ($\partial E/\partial k$).

$$a = \frac{dv_g}{dt} = \frac{\partial v_g}{\partial k} \frac{dk}{dt} = \frac{F}{\hbar} \frac{\partial v_g}{\partial k} = \frac{F}{\hbar^2} \frac{\partial^2 E}{\partial k^2} \quad (31)$$

Again compare to Newton's Law $F = ma$:

$$m \longrightarrow m_e^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2} \right)^{-1} \quad (32)$$

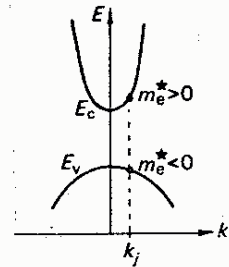
m_e^* is the effective mass of the electron and is a function of the radius of curvature of the E vs. k curve at the given value of k . Large curvature results in small effective mass. Using the effective mass an electron can be treated as a free particle with a reduced mass that accounts for the effects of the lattice potential.

For conduction band:

$$\frac{\partial^2 E}{\partial k^2} > 0 \longrightarrow m_e^* > 0 \quad (33)$$

For valence band

$$\frac{\partial^2 E}{\partial k^2} < 0 \longrightarrow m_e^* < 0!! \quad (34)$$



Due to Bragg reflection the particle accelerates in direction opposite to force. Can be considered equivalent to a particle with a positive mass and positive charge (opposite of electron).

The E vs. k curve is even ($E(k) = E(-k)$) so the electron velocity (derivative) is odd. Therefore for a full band, for every electron with a given velocity there is another electron with the opposite velocity. The conduction band is nearly empty so the current density (current per unit area) is given by

$$J_{cb} = \frac{1}{V} \sum_{CB} (-q)v_i \quad (35)$$

The valence band is nearly full

$$J_{vb} = \frac{1}{V} \sum_{VB} (-q)v_i = \frac{1}{V} \sum_{\text{Filled band}} (-q)v_i - \frac{1}{V} \sum_{\text{Empty states}} (-q)v_i \quad (36)$$

and no current can flow in a full band so

$$J_{vb} = 0 - \frac{1}{V} \sum_{\text{Empty states}} (-q)v_i = \frac{1}{V} \sum_{\text{Empty states}} (q)v_i \quad (37)$$

The behavior of the conduction band can be expressed by treating the vacant states as if they were particles called "holes".

$$m_h^*(k) = -m_e^*(k), \quad q_h = -q_e \quad (38)$$

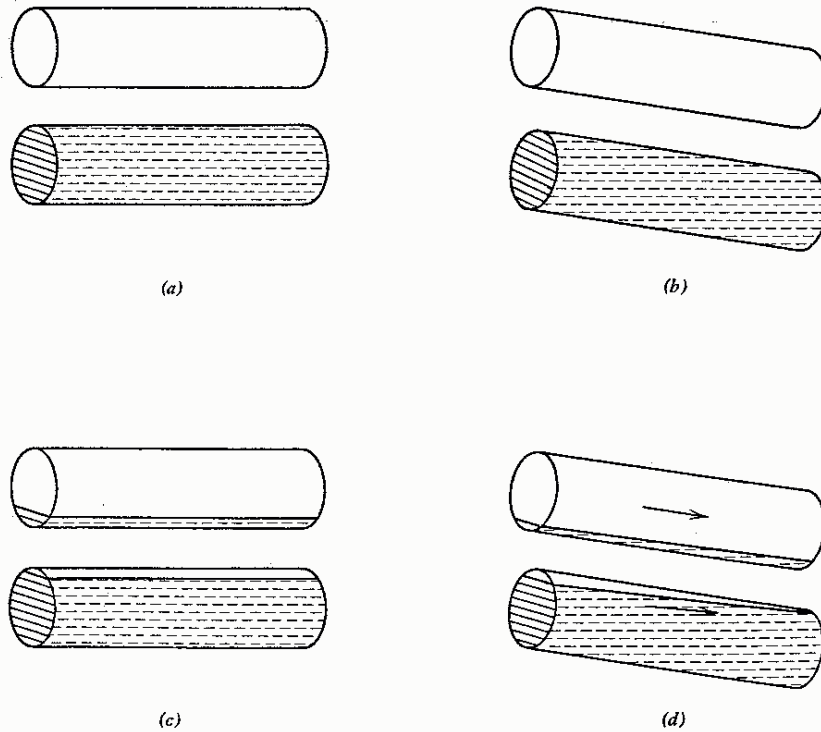
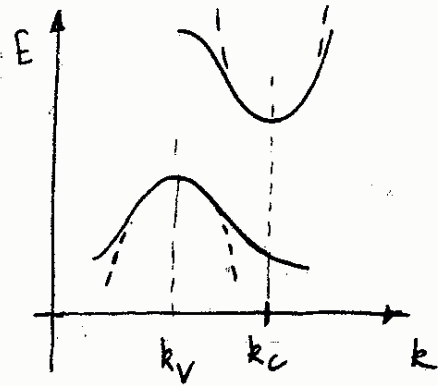


Figure 1.7 Fluid analogy for a semiconductor. (a) and (b) No flow can occur in either the completely filled or completely empty tube. (c) and (d) Fluid can move in both tubes if some of it is transferred from the filled tube to the empty one, leaving unfilled volume in the lower tube.

The hole concept is very useful for valence band. It is easier to describe the behavior of a few empty states than of many full states. In general m^* is not a constant but we will approximate the bands by parabolas (constant curvature) since the full states in the conduction band are all near the bottom while the empty states in the valence band are near the top.



$$E \approx E_C + \frac{\hbar^2 |k - k_c|^2}{2m_e^*} \quad (39)$$

$$E \approx E_V - \frac{\hbar^2 |k - k_v|^2}{2m_h^*} \quad (40)$$

In a real semiconductor the band structures are more complex.

- direct and indirect band gaps (phonon interactions)
 - multiple valence bands with different curvature (heavy and light holes)
 - multiple conduction band minima
 - elliptical rather than spherical minima ($1/m_e^* \approx 1/3(1/m_x + 1/m_y + 1/m_z)$, inverse effective mass actually a tensor, $k^2/m_e^* \rightarrow kM^{-1}k^T$, $m_{ij}^{-1} = \partial^2 E / \partial k_i \partial k_j$)
- $$a = \frac{F}{m} \Rightarrow \vec{a} = M^{-1} \vec{F}$$

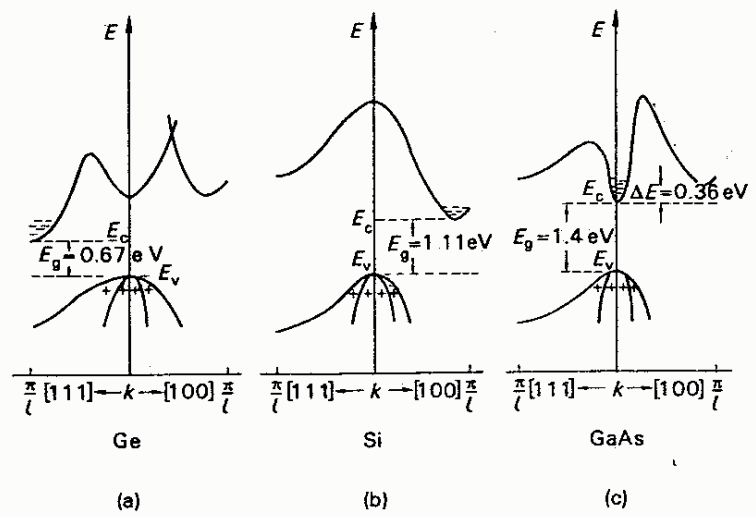
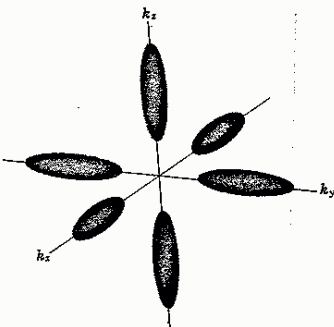


Figure 6.10. The energy bands in momentum space for the important semiconductors, at 300°K: (a) Ge; (b) Si; (c) GaAs.

Intrinsic Carrier Concentration

Carrier Concentration = Density of States \times Occupation Probability

Density of States: Assume parabolic approximation. Rectangular crystal of dimensions L_x, L_y, L_z and atomic dimensions l_x, l_y, l_z . Then there are

$$n_x = \frac{L_x}{l_x}, \quad n_y = \frac{L_y}{l_y}, \quad n_z = \frac{L_z}{l_z}$$

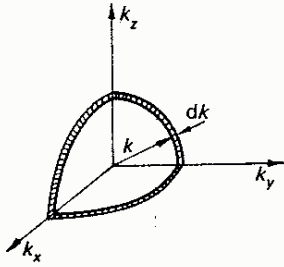
different values of k_x, k_y, k_z per band (states). From reduced zone model, each band has a size equal to a Brillouin zone ($2\pi/l_x, 2\pi/l_y, 2\pi/l_z$). Therefore the spacing of states in k -space is given by

$$\Delta k_x = \frac{2\pi/l_x}{n_x} = \frac{2\pi}{L_x}, \text{ etc.} \quad (41)$$

We could also see this since, for a finite crystal, the solutions must have the periodicity of the whole crystal so $k_x = 2n\pi/L_x$, etc.

Volume per state (1/2 included to account for spin) is given by

$$\frac{1}{2} \Delta k_x \Delta k_y \Delta k_z = \frac{1}{2} \frac{(2\pi)^3}{L_x L_y L_z} \quad (42)$$



The number of allowed states in a spherical shell of radius $|k|$ and thickness dk is the ratio of shell volume to volume per state.

$$dN' = \frac{4\pi k^2 dk}{(2\pi)^3 / 2 L_x L_y L_z} = \left(\frac{|k|}{\pi}\right)^2 L_x L_y L_z dk \quad (43)$$

The density of states per unit crystal volume then is

$$dN = \left(\frac{|k|}{\pi}\right)^2 dk \quad (44)$$

Using the assumed parabolic relation between energy and wavevector, the density states per change in energy is

$$dN = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_C)^{1/2} dE = N_C(E) dE \quad (45)$$

N_C is the density of states in the conduction band ($N_C(E_C) = 0$, $N_C(E) > 0$ for $E > E_C$).

A similar expression can be derived for the valence band.

$$dN(\text{valence}) = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V - E)^{1/2} dE = N_V(E) dE \quad (46)$$

Probability of Occupation

- Each level is a cell into which one or more particles can be placed.
- Calculate number of ways in which n particles can be arranged in those cells using the applicable rules and with total energy fixed.
- Proper distribution maximizes the number W of such arrangements. $S = k \ln W$. $G = E - TS$.

There are three main distributions

1. Classical: any number of distinguishable particles per cell/state. Maxwell-Boltzmann distribution:

$$f_{MB}(E) = \exp\left(-\frac{E}{kT}\right) \quad (47)$$

2. Quantum: indistinguishable particles with symmetric wave functions (bosons, ie. photons, phonons), unlimited number of particles per state. Bose-Einstein distribution:

$$f_{BE}(E) = \frac{1}{\exp\left(\frac{E}{kT}\right) - 1} \quad (48)$$

3. Pauli exclusion principle: indistinguishable particles with antisymmetric wavefunctions (fermions, ie. electrons), up to two particles per state (opposite spins). Fermi-Dirac distribution:

$$f_{FD}(E) = \frac{1}{\exp\left(\frac{E - E_f}{kT}\right) + 1} \quad (49)$$

E_f = Fermi level = maximum energy of occupied state at 0°K . $f_{FD}(E_f) = 1/2$

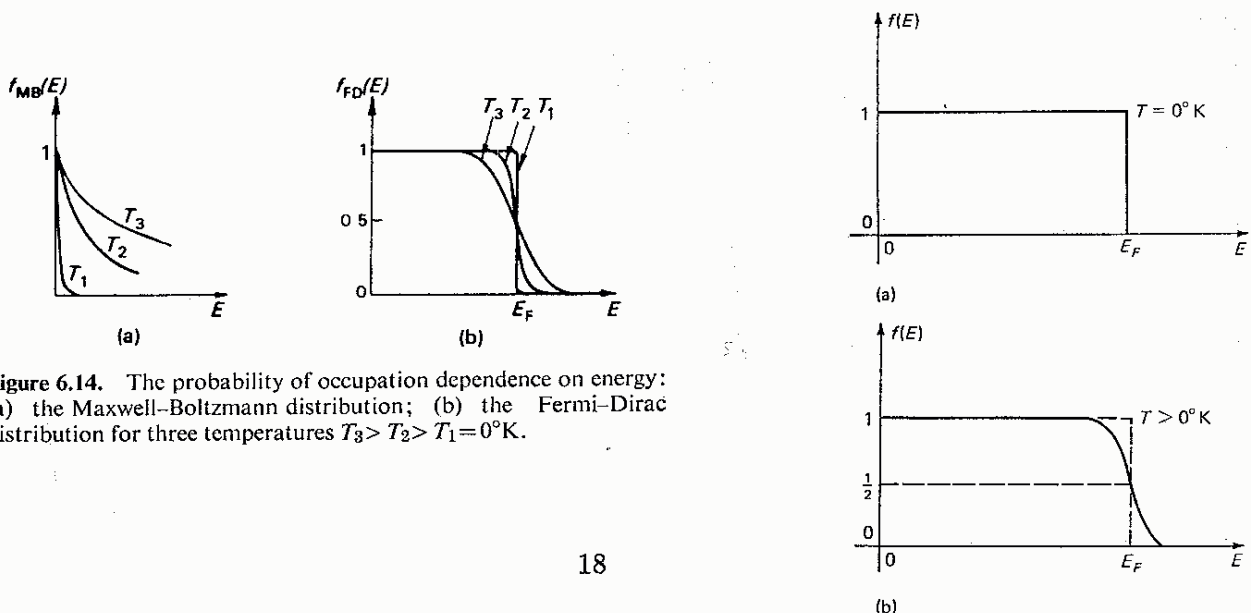


Figure 6.14. The probability of occupation dependence on energy: (a) the Maxwell-Boltzmann distribution; (b) the Fermi-Dirac distribution for three temperatures $T_3 > T_2 > T_1 = 0^\circ\text{K}$.

Two limiting cases (Boltzmann approximation)

$$E - E_f > 4kT : f_{FD} \rightarrow f_{MB}((E - E_f)) = \exp\left(-\frac{E - E_f}{kT}\right) \ll 1 \quad (50)$$

$$E - E_f < -4kT : f_{FD} \rightarrow 1 - f_{MB}((E_f - E)) = 1 - \exp\left(\frac{E - E_f}{kT}\right) \approx 1 \quad (51)$$

To calculate the carrier concentration, the product of the density of available states must be multiplied by the probability of occupation and integrated over the band.

$$n = \int_{E_C}^{\infty} f_{FD}(E) N_C(E) dE \approx \int_{E_C}^{\infty} f_{MB}(E - E_f) N_C(E) dE \quad (52)$$

since for intrinsic material the Fermi level is near the center of the band gap, well below the conduction band minima. Similarly,

$$p = \int_{-\infty}^{E_V} [1 - f_{FD}(E)] N_V(E) dE \approx \int_{-\infty}^{E_V} f_{MB}(E_f - E) N_V(E) dE \quad (53)$$

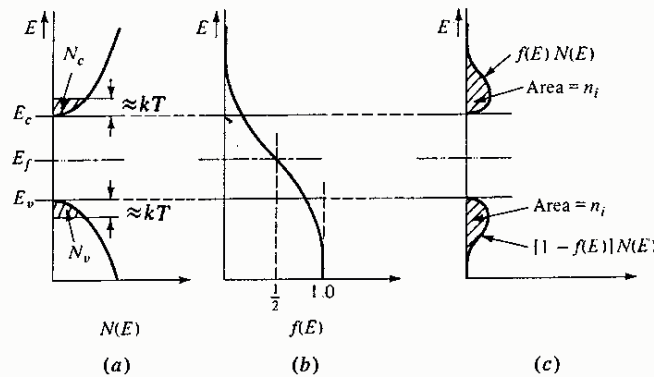
As long as the Boltzmann approximation is valid these integrals can be rewritten as

$$n = N_C \exp\left(-\frac{E_C - E_f}{kT}\right); \quad N_C = 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \quad (54)$$

$$p = N_V \exp\left(-\frac{E_f - E_V}{kT}\right); \quad N_V = 2 \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \quad (55)$$

where N_C and N_V are the effective density of states for the conduction and valence band respectively. They represent the equivalent density of states at the minima of the conduction band or the maxima of the valence band that would give the same electron and hole concentrations as the actual distribution of states in the bands. Equations (54) and (55) are only valid when the Boltzmann approximation is valid.

In silicon $N_C = 2.8 \times 10^{19} \text{cm}^{-3}$ and $N_V = 1 \times 10^{19} \text{cm}^{-3}$ at 25°C .



Conduction Effective Mass vs Density of States Effective Mass

Conduction effective mass is proper effective mass and is used to determine the effect on a carrier of applied forces (e.g., electric field).

For a material with elliptically-symmetric conduction band minima (e.g., silicon, germanium, GaAs):

$$\frac{1}{m_e^*} \cong \frac{1}{3} \left(\frac{1}{m_x} + \frac{1}{m_y} + \frac{1}{m_z} \right) = \frac{1}{3} \left(\frac{1}{m_l} + \frac{2}{m_t} \right) \quad (56)$$

For a material with two spherically-symmetric valence band maxima: (e.g., silicon, germanium, GaAs):

$$m_h^* \cong \frac{N_{hh}m_{hh}^* + N_{lh}m_{lh}^*}{N_{hh} + N_{lh}} = \frac{m_{hh}^{*3/2}m_{hh}^* + m_{lh}^{*3/2}m_{lh}^*}{m_{hh}^{*3/2} + m_{lh}^{*3/2}} \quad (57)$$

The density of states effective mass is a quantity calculated so that the density of states equation derived for materials with only a single spherically-symmetric conduction band minima or valence band maxima can be used for other materials as well.

$$N_c = 2 \left(\frac{2\pi m_e^* kT}{h^2} \right)^{3/2}$$

For example, in silicon where the conduction band has six equivalent minima,

$$m_{e(\text{dos})}^{*3/2} = 6(m_x m_y m_z)^{1/2} = 6(m_t^2 m_l)^{1/2} \quad (58)$$

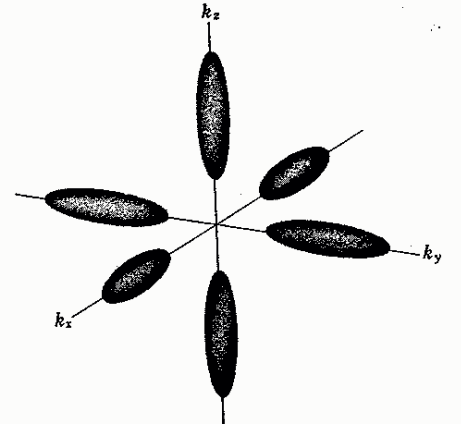
$$m_{e(\text{dos})}^* = 6^{2/3} m_t^{2/3} m_z^{1/3} \quad (59)$$

and the valence band has two different maxima,

$$m_{h(\text{dos})}^{*3/2} = m_{hh}^{*3/2} + m_{lh}^{*3/2} \quad (60)$$

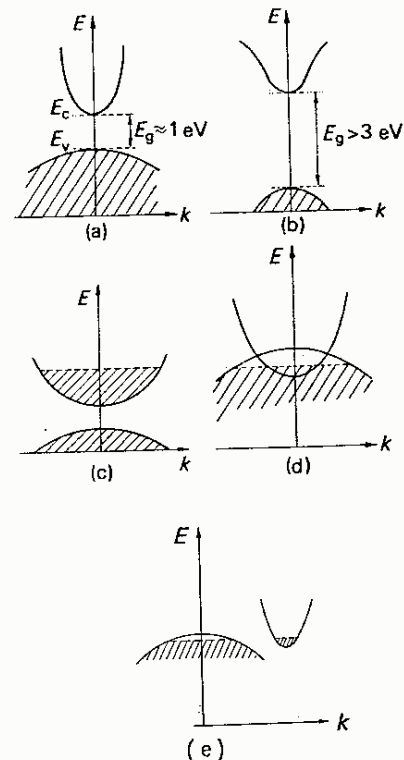
$$m_{h(\text{dos})}^* = \left(m_{hh}^{*3/2} + m_{lh}^{*3/2} \right)^{2/3} \quad (61)$$

Values for silicon, germanium and GaAs are given in text.



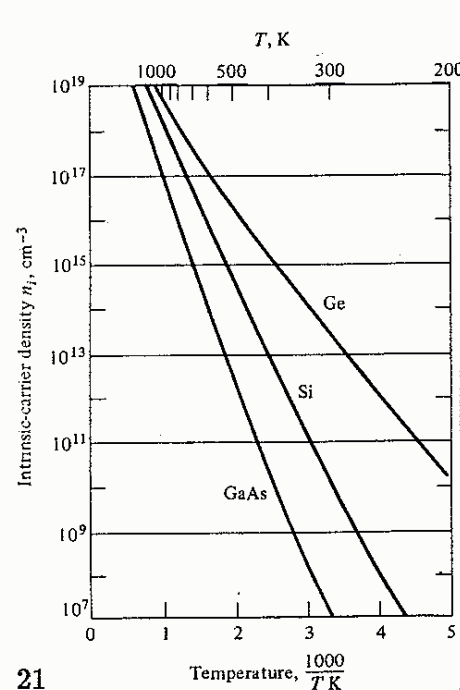
Summary – Energy Bands

- Shape of energy bands depends on periodic potential.
- The energy bands determine the electronic properties of the material.
- Bandgap determines conductivity:
 - No gap (overlapping or partially filled bands) – metal
 - Small gap – semiconductor
 - Large gap – insulator
- Shape of minima/maxima determines electron/hole mobility and density of states
- Direct/indirect bandgap determines light interactions



Number of intrinsic carriers (n_i) as $f(T)$ for 3 common semiconductors. n_i increases as temperature increases and energy gap decreases.

$$E_g = E_{g0} - \beta(T - T_0)$$



Doping in Semiconductors

Multiplying together our expressions for n and p ,

$$np = N_c N_v \exp -\frac{E_c - E_v}{kT} \cong KT^3 \exp -\frac{E_g}{kT} \quad (62)$$

$$K \cong \text{constant}, \quad E_g = 1.12 \text{ eV in Si at } 27^\circ\text{C}$$

- $np = f(f(E), N(E))$ only.
- np independent of impurities.
- np independent of E_f .
- $np \cong 2.1 \times 10^{20} \text{ cm}^{-6}$ for Si at 27°C .

In a pure semiconductor, $p = n = n_i$.

$$n_i = \sqrt{N_c N_v} \exp -\frac{E_g}{2kT} \cong 1.45 \times 10^{10} \text{ cm}^{-3} \text{ for Si at } 27^\circ\text{C}. \quad (63)$$

Also since $n = p$,

$$N_c \exp -\frac{E_c - E_f}{kT} = N_v \exp -\frac{E_f - E_v}{kT} \quad (64)$$

Therefore,

$$\begin{aligned} E_f(n = p) = E_i &= \frac{1}{2} \left(E_c + E_v + kT \ln \frac{N_v}{N_c} \right) \\ &= \frac{1}{2} (E_c + E_v) + \frac{3}{4} kT \ln \frac{m_h^*}{m_e^*} \\ &\cong \frac{1}{2} (E_c + E_v) \end{aligned} \quad (65)$$

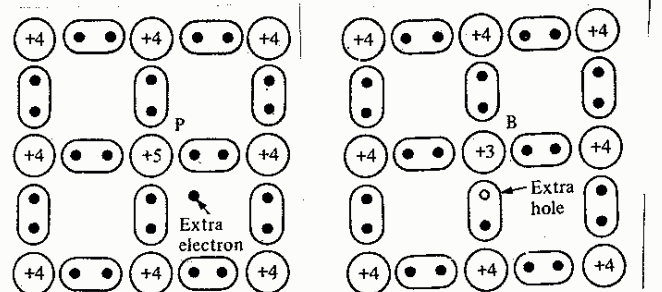
The intrinsic Fermi level is just below midgap.

For both doped and undoped semiconductors:

$$n = n_i \exp \frac{E_f - E_i}{kT} \quad (66)$$

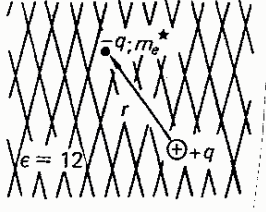
$$p = n_i \exp \frac{E_i - E_f}{kT} \quad (67)$$

To increase the conductivity of a pure (intrinsic) semiconductor, it is possible to add impurity atoms to the lattice that will donate (valence 5, ie, P, As, Sb) or accept (valence 3, ie, B) electrons resulting in the creation of additional conduction electrons or holes, respectively.



Dopant atoms occupy substitutional lattice sites and the extra electron or hole is very loosely bound. In other words they can easily move to the conduction and valence bands, respectively, with only small thermal excitations.

Imagine a Bohr-type atom (like hydrogen atom). Balancing centripetal force and electric attraction:



$$\frac{q^2}{4\pi\epsilon r^2} = \frac{m_e^* v^2}{r}.$$

For the lowest energy level, the DeBroglie wavelength is given by:

$$2\pi r = \lambda = \frac{h}{p} = \frac{h}{m_e^* v}$$

$$v = \frac{h}{2\pi m_e^* r}.$$

Substituting,

$$r = \frac{h^2 \epsilon}{q^2 m_e^* \pi} = r_H \left(\frac{\epsilon}{\epsilon_0} \right) \left(\frac{m_0}{m_e^*} \right), \quad (68)$$

where r_H is the radius of the electron orbit in a hydrogen atom and ϵ_0 is the permittivity of free space. Therefore, the radius of the electron orbit is about 50 times that of a hydrogen atom and the extra electron ranges over a large portion of the lattice, which has a atomic spacing of less than 5 times the atomic radius of hydrogen.

The energy of that electron is given by the sum of the kinetic and potential energy.

$$\begin{aligned} E &= \frac{m_e^* v^2}{2} - \frac{q^2}{4\pi\epsilon r} \\ &= -\frac{q^2}{8\pi\epsilon_0 r_H} \left(\frac{\epsilon_0}{\epsilon} \right)^2 \left(\frac{m_0^*}{m_0} \right) \\ &= E_H \left(\frac{\epsilon_0}{\epsilon} \right)^2 \left(\frac{m_0^*}{m_0} \right) \end{aligned} \quad (69)$$

where E_H is -13.6 eV. Therefore, for silicon the ionization energy for an impurity is approximately 0.025 eV, while the average electron thermal energy at room temperature is about 0.026 eV. As we will see later, at room temperature virtually all shallow donors and acceptors are ionized.

For every free electron or hole which results from the ionization of a substitutional impurity, an oppositely charged impurity ion also results. The result is the charge neutrality condition.

$$\underbrace{n + N_a^-}_{-} \stackrel{\approx}{=} \underbrace{p + N_d^+}_{+} \quad (70)$$

Assuming that all impurities are ionized,

$$p - n \stackrel{\approx}{=} N_a - N_d \quad (71)$$

Combining this expression with the law of mass action ($np = n_i^2$), the hole and electron concentration can be calculated as a function of the net doping, $N_d - N_a$. For material heavily doped enough to be extrinsic ($|N_d - N_a| \gg n_i$) with only a single dopant type, the minority carrier concentration can be neglected relative to the majority carrier concentration:

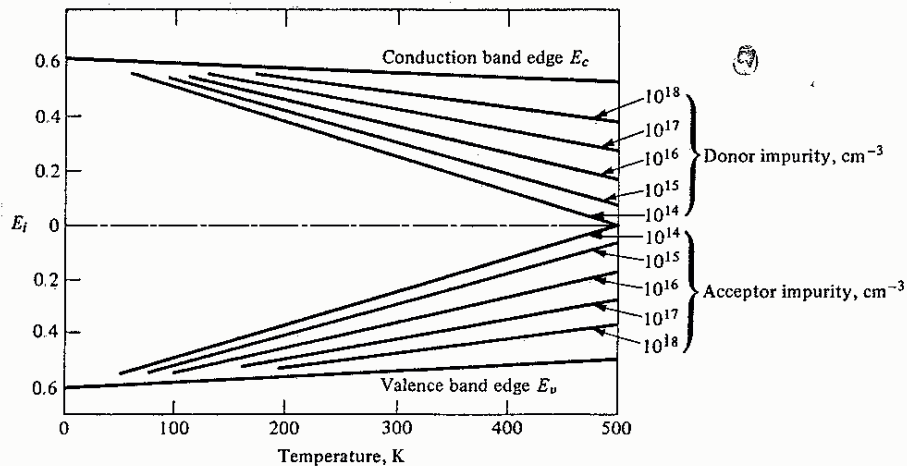
$$\overbrace{n_n \cong N_d - N_a; \quad p_n = n_i^2 / (N_d - N_a)}^{\text{n-type}} \quad (72)$$

$$\overbrace{p_p \cong N_a - N_d; \quad n_p = n_i^2 / (N_a - N_d)}^{\text{p-type}} \quad (73)$$

But for very lightly-doped material or at high temperatures, the mass-action and charge-neutrality equations must be solved simultaneously to find the carrier concentrations. The Fermi level can be found based on knowledge of the carrier concentration using Equations (38) and (39)

$$E_f = E_c - kT \ln \frac{N_c}{n} = E_v + kT \ln \frac{N_v}{p} \quad (74)$$

As temperature increases, n_i increases and so n and p tend to become more nearly equal. In addition, the energy gap also narrows slowly with temperature.



Summary – Shallow Dopants at Equilibrium

Assumptions:

1. Complete impurity ionization ($N_d^+ = N_d$, $N_a^- = N_a$)
2. Thermal equilibrium

Equations:

Charge neutrality:
$$n - p + N_a^- - N_d^+ \cong 0$$

Mass Action (Equilibrium):
$$np = n_i^2$$

Fermi-Dirac Distribution:
$$f(E) = \frac{1}{1 + \exp[(E - E_f)/kT]}$$

Maxwell-Boltzmann Approximation:
$$f(E) \cong \exp\left(-\frac{E - E_f}{kT}\right) \quad (\text{valid for } E - E_f \geq 3kT)$$

$$f(E) \cong 1 - \exp\left(\frac{E - E_f}{kT}\right) \quad (\text{valid for } E - E_f \leq -3kT)$$

Carrier concentrations:
$$n = n_i \exp \frac{E_f - E_i}{kT} = N_c \exp -\frac{E_c - E_f}{kT}$$

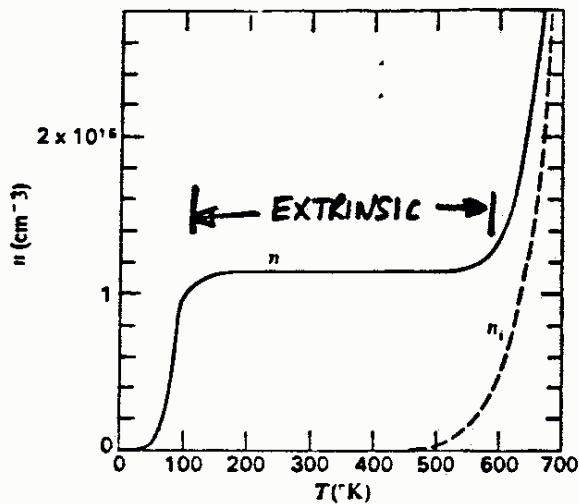
$$p = n_i \exp \frac{E_i - E_f}{kT} = N_v \exp -\frac{E_f - E_v}{kT}$$

Extrinsic Semiconductors:
$$n_n \cong N_d^+ - N_a^- \qquad n_p \cong \frac{n_i^2}{(N_a^- - N_d^+)}$$

$$p_n \cong \frac{n_i^2}{N_d^+ - N_a^-} \qquad p_p \cong (N_a^- - N_d^+)$$

$$\underbrace{E_f - E_c \cong -kT \ln \frac{N_c}{N_d^+ - N_a^-}}_{n\text{-type}} \quad \underbrace{E_f - E_v \cong kT \ln \frac{N_v}{N_a^- - N_d^+}}_{p\text{-type}}$$

Low Temperature



Electron concentration as $f(T)$ for Si doped with $N_d \cong 1 \times 10^{16}/\text{cm}^3$. Normal operation of semiconductor device is in the extrinsic conductivity range where purposely introduced impurities control the electrical properties.

At low temperatures, freezeout of the impurities occurs (insufficient thermal energy to excite electrons the $\approx 0.05\text{eV}$ from E_d to E_c) and

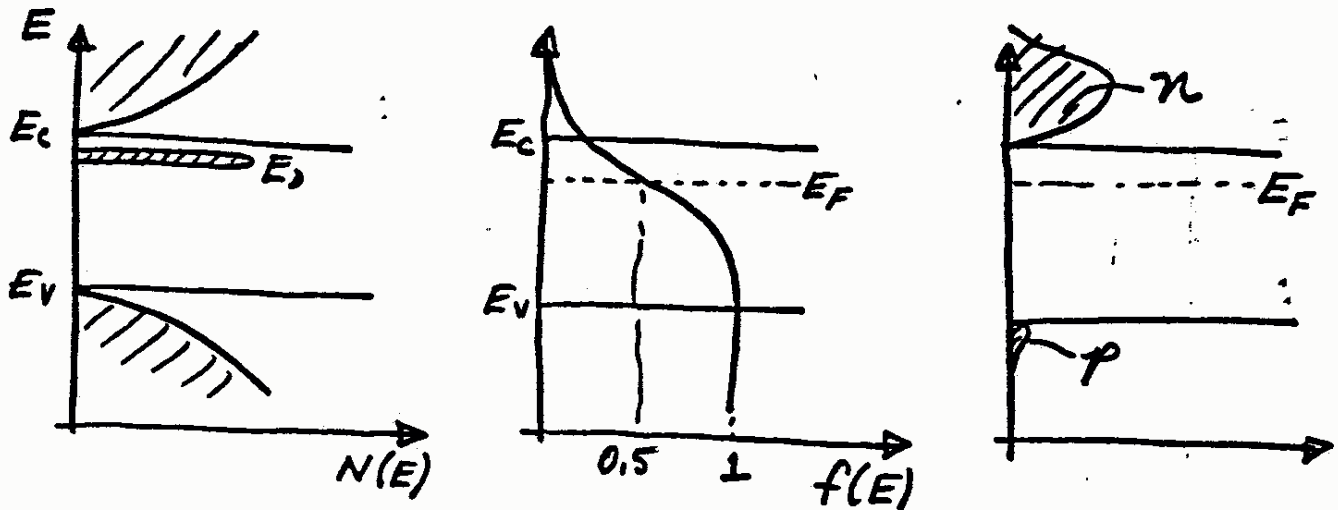
$$N_d^+ = N_d - \frac{N_d}{1 + (1/2) \exp(E_d - E_f)/kT}$$

↑ 2 spins of electron in donor level

$$N_a^- = \frac{N_a}{1 + 1/4 \exp(E_a - E_f)/kT}$$

↑ 2 spins × 2 valence bands in Si, Ge, GaAs

Degeneracy



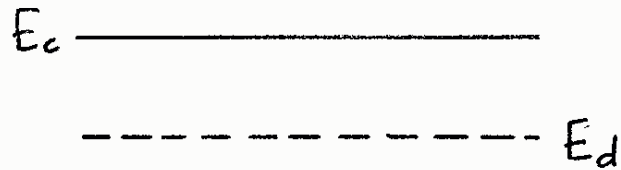
1. E_f must move up so $f(E)$ will increase in the conduction band as to increase the probability of e^- being there.
2. When E_f increases, $1 - f(E)$ decreases in the valence band. Therefore, p decreases.
3. $f(E)$ at E_d must be low if these states are to all be ionized (i.e., $E_f < E_d$).

This implies that $E_f \cong E_d$ or in CB \Rightarrow degenerate state.

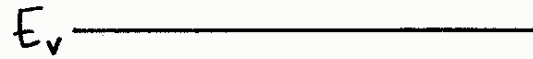
This happens at doping levels comparable to effective densities of states N_c and N_v . N_d or $N_a \sim 10^{19}/\text{cm}^3$.

Deep Donors or Acceptors

Suppose we add an impurity whose donor level is not very close to the conduction band ($E_c - E_f > 4kT$). What are n , p , E_f ?



Cannot assume $n = N_d$ in this case. Thermal energy may not be enough to ionize all impurities.



$\Pr\{e^- \text{ occupies donor state}\}$ (not ionized) is given by:

$$E_f < E_d - 3kT : \quad f(E) \cong \exp\left(-\frac{E_d - E_f}{kT}\right)$$

$$E_f > E_d + 3kT : \quad 1 - f(E) \cong \exp\left(-\frac{E_f - E_d}{kT}\right)$$

Therefore, the number of ionized donors is:

$$E_f < E_d - 3kT : \quad N_d^+ \cong N_d \underbrace{\left[1 - \exp\left(-\frac{E_d - E_f}{kT}\right)\right]}_{\Pr\{e^- \text{ is not there}\}}$$

$$E_f > E_d + 3kT : \quad N_d^+ \cong N_d \exp\left(-\frac{E_f - E_d}{kT}\right)$$

(Exact expressions in class notes page 26.)

Neutrality: $\underbrace{n}_{-} = \underbrace{p + N_d^+}_{+} \leftarrow \text{ionized donors}$

How do we calculate E_f ? Substitute expressions in neutrality condition.

$$N_c \exp\left(\frac{E_f - E_c}{kT}\right) = N_d \exp\left(\frac{E_d - E_f}{kT}\right) \quad (\text{assume } p \ll n)$$

But,

$$E_d - E_f = \underbrace{(E_d - E_c)}_{\text{known}} - (E_f - E_c)$$

Therefore, the only unknown is $(E_f - E_c)$ which can be solved for. Once E_f is known, then (assuming $N_d^+ \gg n_i$):

$$n \cong N_d^+ \quad \text{and} \quad p \cong \frac{n_i^2}{N_d^+}$$

Some specific examples

Example 1.

Suppose we have a silicon bar with both donor and acceptor impurities (shallow) present.
For example,

$$10^{15}/\text{cm}^3 \text{ Arsenic}$$

$$5 \times 10^{16}/\text{cm}^3 \text{ Boron}$$

both shallow \Rightarrow complete ionization

Thus,

$$p \cong N_a - N_d = 4.9 \times 10^{16}/\text{cm}^3$$

$$n \cong \frac{n_i^2}{N_a - N_d} \cong 4.6 \times 10^3/\text{cm}^3$$

$$E_f \cong E_v + kT \ln \frac{N_v}{N_a - N_d}$$

$$\cong 0.133 \text{ eV above valence band.}$$

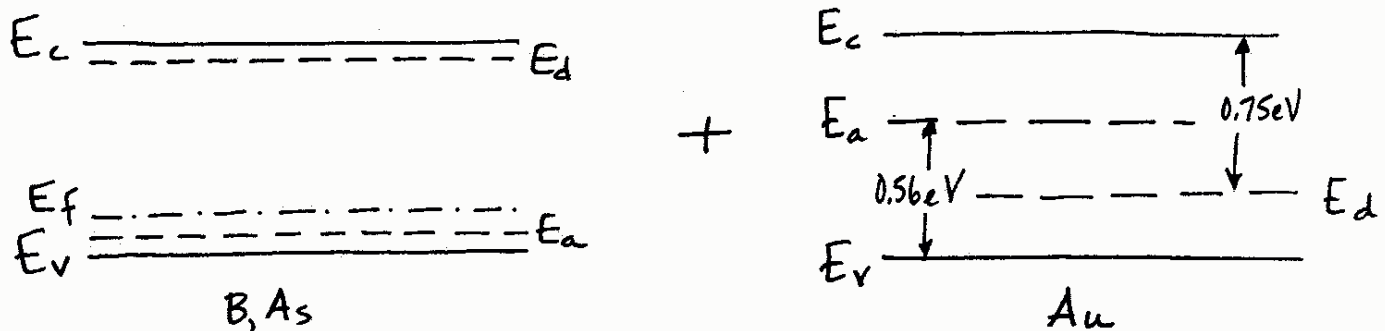
$$E_c \text{ --- } E_d$$

Therefore, p dopant dominates and
 E_f is near the valence band.

$$E_f \text{ --- } E_a$$

Example 2.

Suppose we now add to the above silicon an impurity which adds deep level states within the band gap.



If the Au concentration is $1 \times 10^{15}/\text{cm}^3$. What are n , p , E_f ?

Since $[\text{Au}] \ll [\text{B}] - [\text{As}]$, assume E_f approximately the same.

Therefore,

1. Au donor states will be ionized since $E_f < E_d$ and, therefore, $f(E) \approx 0$ and e^- will not occupy donor states.
2. Au acceptor states will not be ionized since $E_f < E_a$ and, therefore, $1 - f(E) \cong 1$ and e^- will not occupy acceptor states.

Therefore,

$$p \cong N_{a_B} - N_{d_{As}} - N_{d_{Au}} \cong 4.8 \times 10^{16}/\text{cm}^3$$

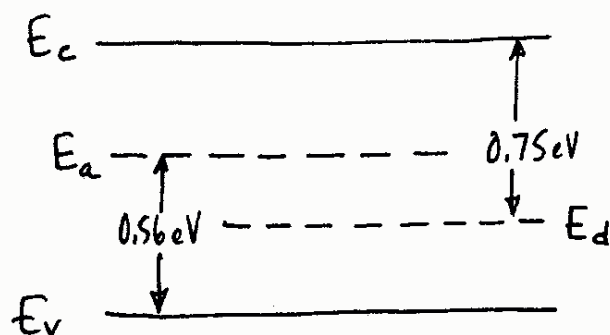
$$n \cong \frac{n_i^2}{p} \cong 4.7 \times 10^3/\text{cm}^3$$

$$E_f - E_v = kT \ln \frac{N_v}{N_{a_B} - N_{d_{As}} - N_{d_{Au}}} \\ \cong 0.134 \text{ eV above valence band}$$

Thus, the net carrier concentrations and E_f have been only slightly changed. The Au has slightly decreased p . Therefore, p increases very slightly.

Example 3.

Suppose we had only the deep lying impurity Au in otherwise intrinsic Si.



It is reasonable to expect that some Au atoms will act as acceptors and some as donors. Since levels are deep, n and p will be small (but likely the dominant will be greater than n_i).

Also, probably the silicon will become p type because E_a is closer to E_v than E_d is to E_c (e.g., easier to ionize h^+ than e^-).

Let N = number of Au atoms/cm³.

The number of ionized donors is given by

$$N_d^+ = (N - N_a^-) \underbrace{\left[1 - \frac{1}{1 + \exp(E_d - E_f/kT)} \right]}_{\text{Pr}\{e^- \text{ is not there}\}}$$

The number of ionized acceptors is given by

$$N_a^- = (N - N_d^+) \underbrace{\left[\frac{1}{1 + \exp(E_a - E_f/kT)} \right]}_{\text{Pr}\{e^- \text{ is there}\}}$$

(These expressions use F-D statistics. Once the Fermi level is approximately located, simpler M-B expressions could be used.)

We also note that

$$N^* + N_a^- + N_d^+ = N$$

where N^* = neutral Au atoms.

Also,

$$\begin{aligned} n &= N_c \exp - \frac{E_c - E_f}{kT} \\ p &= N_v \exp - \frac{E_f - E_v}{kT} \\ n + N_a^- &= p + N_d^+ \end{aligned}$$

This gives us 6 equations with 6 unknowns ($n, p, E_f, N_a^-, N_d^+, N^*$). Therefore, there is a unique solution although it is not trivial to obtain! In many instances the problem can be greatly simplified by making the appropriate assumptions about the location of the Fermi level.