Lecture 15. Electronic Properties of Semiconductors

- Intrinsic and doped semiconductors
- Doping methods
- Law of Mass Action
- Chemical potential in semiconductors
- Conductivity of intrinsic and doped semiconductors
Intrinsic and Doped Semiconductors

The band gap $E_g \lesssim 3 \, eV$. Electrons and holes can be introduced by several processes: thermal excitations, photon absorption, field effect, etc.

**Intrinsic semiconductors** contain a negligibly small amount of impurities compared with thermally generated electrons and holes.

**Doped semiconductors**: doping enables an orders-of-magnitude increase of the mobile charge carrier concentration, due to significant reduction of the activation energies (separately for electrons and holes).

Controllable doping starts with fabrication of extremely pure intrinsic semiconductors. Unintentional doping concentration has been reduced below $\sim 10^{16} \, m^{-3}$, which is comparable to the concentration of thermally-excited carriers in intrinsic semiconductors at RT. The atom concentration in Si is $\sim 5 \times 10^{28} \, m^{-3}$, so the material purity is in the part-per-trillion range!

Moderate doping preserves the band structure of the host material. It provides the charge carriers that otherwise wouldn’t be available. However, doping also produces disorder (carrier scattering by charged donors). In some 2D cases, e.g. in GaAs/AlGaAs heterostructures, carriers and parent charged donors can be spatially separated.
Typical semiconductors (such as Si and Ge) have valence electrons in \( ns \) and \( np \) orbitals \((n = 3, 4, \ldots)\). These orbitals form 8 (bonding and anti-bonding) bands. There are eight valent electrons from two atoms in each unit cell. As a result, four low-energy bonding bands are fully occupied, while the anti-bonding bands are empty.

Spin-orbit coupling further splits the three \( p \) orbitals – as a result, there are several bands with “light” and “heavy” holes. In most semiconductors, the electron band structure and dynamics are simpler than those for holes.

The band structure for GaAs near the Brillouin zone center \( \Gamma \). The conduction band is an \( s \) band, while the three valence bands closest to the Fermi energy are \( p \) bands.
Typical semiconductors – *Si*, *Ge*, *GaAs* – have several valence and conduction bands.

For these materials, the valence bands are similar (though the effective hole mass is different for different valence bands). The maxima of all valence bands are at the BZ center.

The conduction bands are quite different. The minimum of the conduction band in GaAs is also at the BZ center (*direct bandgap* semiconductor), for Si and Ge – at a non-zero \( k \) (*indirect bandgap* semiconductors).
Even with heavy doping, the carrier concentration rarely exceeds $10^{26} m^{-3}$. The atom concentration in Si is $\sim 5 \times 10^{28} m^{-3}$. Thus, electrons occupy only a very small fraction of the conduction band where their energy is lowest (and holes occupy only a very small fraction of the valence bands). We are interested in the properties of the energy spectrum in the vicinity of the $\varepsilon(k)$ minimum for the conduction band and in the vicinity of its maximum for the valence band.

Direct-bandgap $\text{Ge}$

Indirect-bandgap $\text{Si}$
Effective Masses in bulk Si, Ge, and GaAs

The effective mass is a tensor.

Averaging is different for the effective masses that are relevant to the conductivity and DoS.

When the electron motion is restricted to 2D, the effective masses are different: in [100] n-type Si MOSFETs $m^*_{e,\sigma} \approx 0.2 m_e$
**Shallow impurities**: large-radius Bohr’s orbit and small binding energy (singly-ionized ions + small effective mass + large dielectric constant). Energy is either close to the top of the VB, or to the bottom of the CB. These impurities are most important for applications.

**Deep traps**: small-radius orbit and large binding energy (e.g., multiply-ionized atoms).
Shallow Impurities (Dopants)

Shallow dopants (donors and acceptors) can be viewed as Hydrogen-like “atoms” with an effective electron mass $m^*$ in a medium with the effective dielectric constant $\varepsilon_r$. Modified electron energy and Bohr’s radius are:

$$E_0 = \frac{m_e e^4}{8\varepsilon_0^2 \hbar^2} = 13.6eV \quad a_0 = \frac{4\pi \varepsilon_0 \hbar^2}{m_e e^2} \approx 0.05nm$$

$H$ atom

$$E_{eff} = E_0 \left( \frac{1}{\varepsilon_r^2} \frac{m^*}{m_e} \right) \quad a_{0eff} = a_0 \left( \frac{\varepsilon_r m_e}{m^*} \right)$$

For donors in Si: $m^* \approx 0.2 \, m_e, \varepsilon_r \approx 12 \Rightarrow E_{eff} \approx 20meV, \, a_{0eff} \approx 3nm \gg a$

Thus, the energy of electron “bonding” to donors is much smaller than $E_g$ in Si. The electron states are by $E_{eff}$ lower than the bottom of the conduction band, whereas the hole states are by $E_{eff}$ higher than the top of the valence band.

$k_B 300K = 26meV$

Ionization energies (in meV) of donors and acceptors in Si. Not to scale.

Because $E_D (E_A)$ is so small for shallow dopants, they are ionized at $RT (= 26meV)$ and provide free charge carriers in the conduction or valence bands.

Not to scale!
Doping Techniques

**Bulk Doping**: dopants can be added during the crystal growth process. Used to fabricate wafers for electronic industry.

**Surface doping**: by diffusion and implantation (dominant over the last 20 years for fabrication of nanoelectronic devices).

In electronic industry, the doping requirements span several orders of magnitudes in energy and dose. For example, heavy doping is used to form the source/drain contacts in MOSFETs (next lecture).

Fabrication of modern microprocessors requires *up to twenty* different types of doping per wafer.
Ion Implantation

The ion energies are within the range $10 - 500 \text{ keV}$. A mono-energetic ion beam forms a broad depth distribution of dopants. Under typical conditions the depth ranges between $10\text{nm}$ and $1\mu\text{m}$ (the depth increases with ion energy).

Crystal damage created by ion implantation must be removed by **rapid annealing**. Annealing – to heal the crystal and eliminate dangling bonds (carrier traps), rapid - to prevent dopant diffusion.
Advantages of Ion Implantation

- Precise control of dose and depth profile. Energy of ions controls the depth profile, the ion beam current controls the dose.

- Low-temperature process - photoresist can be used as a self-aligned mask.

- Less sensitive to surface cleaning procedures.

Application example: self-aligned MOSFET source/drain regions

- As$^+$
- Poly Si Gate
- n$^+$
- SiO$_2$
Substitutional doping of semiconductors up to the alloying range is hampered by the solubility of the doping impurity into the host lattice. Above the solubility limit the solid solution becomes thermodynamically unstable, leading to phase separation and formation of aggregates. Among the various out-of-equilibrium techniques developed to overcome this limitation, laser-assisted processes are well adapted to silicon, germanium and their alloys. A laser burst melts the surface of the crystal during a few tens of nanoseconds. After the burst, the speed of re-crystallization is of the order of $10^{10}$ K s$^{-1}$, slow enough to allow crystal reconstruction on the underlying non-melted template, but fast enough to prevent impurity diffusion and precipitation. Doping atoms can be introduced by pre-implantation (laser thermal anneal) or in situ by exposure of the surface to impurity-carrying gas molecules in a vacuum chamber (gas-immersion laser doping or GILD), as illustrated below.

Recently, the GILD technique\textsuperscript{37} has been applied to induce superconductivity in silicon by boron doping\textsuperscript{4}. \textit{In-situ} real-time monitoring of the transient reflectivity at a 675-nm wavelength made it possible to adjust the power and duration of the pulses for optimal doping and to ensure that the dopant incorporation profile was flat with an abrupt interface. High-resolution X-ray diffraction measurements and secondary ion mass spectroscopy on the doped layers demonstrated a maximum doping level of the order of 10 at\% atomic (that is, $\sim 5 \times 10^{21}$ atoms cm$^{-3}$) for typically 500 subsequent laser shots. This substitutional concentration is far above the critical threshold for the MIT (a few $10^{18}$ at cm$^{-3}$) and even above the solubility limit ($\sim 10^{20}$ at cm$^{-3}$). Unfortunately, this technique, well suited to silicon and germanium, cannot be applied to most of the other semiconducting compounds, including diamond.
Impurity Bands, Degenerate Semiconductors

**Inter-dopant distance**

\[ \Delta x \approx \frac{1}{n_d^{1/3}} \]

\[ n_d \sim 10^{22} \text{ m}^{-3} \quad \Delta x \approx 46 \text{ nm} \]

\[ n_d \sim 10^{25} \text{ m}^{-3} \quad \Delta x \approx 5 \text{ nm} \]

In Si, at \( N_{A/D} > 10^{24} \text{ m}^{-3} \) “impurity bands” are formed (separated from CB and VB). This is a misleading term as the wavefunctions may remain localized. Indeed, though the impurity bands are no more than half-filled, they don’t necessarily support extended electron/hole states (disorder \( \Rightarrow \) Anderson MIT, el.-el. correlations \( \Rightarrow \) Mott MIT).

A semiconductor with delocalized states in donor/acceptor band is referred to as “degenerate” (c.f. “degenerate Fermi gas”). For electrons in Si: B the metal-insulator transition in the donor band occurs at \( n_d \approx 4 \times 10^{24} \text{ m}^{-3} \) (the atom concentration in Si is \( \sim 5 \times 10^{28} \text{ m}^{-3} \), so the donor concentration at the MIT is only 0.01%).
(b) At higher concentration, when the inter-impurity distance is small, quantum tunneling tends to delocalize the electron/hole wavefunctions, leading to a metallic behavior at $T = 0$ with a Fermi level pinned inside the impurity band.

(c) Because of strong on-site correlations, the spin-degenerate half-filled impurity band can split into an empty band and a full band (Mott insulator).

(d),(e) On further doping, these two bands begin to overlap and a degenerate semiconductor is formed (the metal–insulator transition takes place).
“Electrons” and “holes” in semiconductors are different from the “electron-like” and “hole-like” quasiparticles in metals.

In metals, the “hole” has no simple “image”. Each band can have both “electrons” and “holes”.

In semiconductors, such an image does exist – excitations are linked to broken covalent bonds. Electrons occupy the conduction band, holes – the valent band. The number of “electrons” and “holes” is constant at a given $T$ - there is a dynamic equilibrium between the number of broken covalent bonds and the number of recombination processes.

The difference is due to the lattice bonding in these materials:
- Metals: metallic bonding;
- Semiconductors: covalent bonding.
Recall: Chemical Potential in Metals

By definition, the chemical potential is the total energy change (kinetic + potential) in the process of adding one particle to a system.

In metals at $T = 0$, adding one electron results in an increase of the kinetic energy of the degenerate electron gas by $E_F$ (potential energy is ignored). At low $T$, the situation is similar - on average, adding an electron increases the system energy by $E_F$. At high (comparable to $\mu$) temperatures:

$$\mu(T) \approx \mu(0) \left[ 1 - \frac{\pi^2}{12} \left( \frac{k_B T}{\mu(0)} \right)^2 \right]$$

- this means that since the occupancy becomes $<1/2$ at $\mu(0)$, the price to pay for adding an electron on average is less than $\mu(0)$. $\mu$ becomes negative when the occupancy at $\epsilon = 0$ becomes less than $\frac{1}{2}$ (a non-degenerate Fermi gas).

The non-degenerate scenario is never realized in metals. For this reason, in metals we used interchangeably the chemical potential and the Fermi energy, and $E_F$ separated filled and empty electron states at $T = 0$.

In semiconductors $\mu$ is usually in the band gap, and its meaning differs from $E_F$ in the “metallic” sense.
Chemical Potential in Intrinsic Semiconductors

In *intrinsic* \((n = p, \text{ no dopants})\) semiconductors, only the “Boltzmann” tails of the Fermi function are in the \(CB\) and \(VB\), there are no electron states near the chemical potential which is in the band gap.

\[
\mu \approx e^{-\beta(E_C - \mu)} + 1
\]

\[
\mu = \frac{E_C + E_V}{2}
\]

- the chemical potential for intrinsic semiconductors is approximately in the middle of the band gap.

If \(\mu\) is close to the middle of the band gap, the Boltzmann approximation is justified, and the electrons and holes form a non-degenerate Fermi gas

\[
|\mu - \varepsilon| \gg k_B T: \text{ e.g., for intrinsic Si } \quad \frac{1}{2}(E_C - E_V) \approx 0.55eV \approx 21 \times k_B 300K.
\]
Law of Mass Action

More accurate approach \((m_e \neq m_h)\)

\[
\begin{align*}
\mathcal{E}_n(T) & \approx \int_{E_C}^{\infty} g_c(\varepsilon) e^{-\beta(\varepsilon - \mu)} d\varepsilon = \frac{m_e \sqrt{2m_e}}{\pi^2 \hbar^3} \int_{E_C}^{\infty} \sqrt{\varepsilon - E_C} e^{-\beta(\varepsilon - \mu)} d\varepsilon = \frac{1}{4} \left( \frac{2m_e k_B T}{\pi \hbar^2} \right)^{3/2} e^{-\beta(E_C - \mu)} \\
n(T) &= \int_{E_C}^{E_V} g_c(\varepsilon) e^{-\beta(\varepsilon - \mu)} d\varepsilon = N_c e^{-\beta(E_C - \mu)} \\
p(T) &= \int_{-\infty}^{E_V} g_v(\varepsilon) e^{-\beta(\mu - \varepsilon)} d\varepsilon = N_v e^{-\beta(\mu - E_V)} \\
\mathcal{E}_n(T) p(T) &= \frac{1}{2} \left( \frac{k_B T}{\pi \hbar^2} \right)^3 (m_e^* m_h^*)^{3/2} e^{-\beta E_g} \equiv n_i^2
\end{align*}
\]

Intrinsic \((i)\)

<table>
<thead>
<tr>
<th>(E_g [\text{eV}])</th>
<th>(n_i [\text{cm}^{-3}])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ge</td>
<td>0.67</td>
</tr>
<tr>
<td>Si</td>
<td>1.1</td>
</tr>
<tr>
<td>GaAs</td>
<td>1.43</td>
</tr>
</tbody>
</table>
Example: Intrinsic Silicon

\[
n_i(T)p_i(T) = \frac{1}{2} \left( \frac{k_B T}{\pi \hbar^2} \right)^3 (m_e^* m_h^*)^{3/2} e^{-\beta E_g} = \equiv N_C N_V e^{-\beta E_g}
\]

At home: estimate \( E_g \) from the plot.
Does it match the expected \( E_g \approx 1.1 eV \)?

\[
N_C \approx 2.8 \times 10^{25} \left( \frac{T}{300 K} \right)^{1.5} m^{-3} \quad N_V \approx 1 \times 10^{25} \left( \frac{T}{300 K} \right)^{1.5} m^{-3}
\]

At 300K \( e^{-\beta E_g} = e^{-1.1/0.026} \approx 4 \times 10^{-19} \)

\[
n_i(300K)p_i(300K) \approx 1 \times 10^{32} m^{-6}
\]

\[
n_i(300K) = p_i(300K) \approx 1 \times 10^{16} m^{-3}
\]

This provides an order-of-magnitude estimate for a “reasonable” purity of intrinsic Si: the concentration of dopants should not exceed \( 10^{15} m^{-3} \) to be negligible at \( T = 300K \).

F. Morin and J. Maita.
Phys. Rev. 96, 28 (1954)
\[ \nu \text{ in Intrinsic Semiconductors } (n_i = p_i) \]

\[ n_i(T) = p_i(T) = \frac{1}{\sqrt{2}} \left( \frac{k_B T}{\pi \hbar^2} \right)^{3/2} (m_e^* m_h^*)^{3/4} e^{-\beta E_g/2} \]

\[ n_i(T)p_i(T) = N_C N_V e^{-\beta E_g} = N_C N_V e^{-\beta (E_C - E_V)} \]

\[ 1 = \frac{N_C}{N_V} e^{-\beta (E_V + E_C - 2\mu)} \]

\[ : p_i^2(T) = N_V^2 e^{-2\beta (\mu - E_V)} \]

\[ \mu = \frac{1}{2} (E_V + E_C) + \frac{1}{2} (k_B T) \ln \left( \frac{N_V}{N_C} \right) \]

\[ = \frac{1}{2} (E_V + E_C) + \frac{3}{4} (k_B T) \ln \left( \frac{m_h^*}{m_e^*} \right) \]

\[ \mu \text{ is precisely in the middle of the band gap if } T = 0 \text{ or } m_h^* = m_e^* (N_C = N_V). \]

**Example.** How close to the mid-gap is the chemical potential in Si at \( T = 300K \) (for DoS calculations: \( m_e^* \approx 1.08m_e \), \( m_{lh}^* \approx 0.81m_e \))

\[ \mu = \frac{E_g}{2} + \frac{3}{4} (k_B T) \ln \left( \frac{0.81}{1.08} \right) = 0.55eV - 0.75 \times 0.026eV \times (0.29) = 0.550eV - 0.006eV \]
Carrier Concentration in Doped Semiconductors

Importance of the Law of Mass Action: \( np = n_i^2 \) still holds in doped semiconductors provided that \( \mu \) is in the band gap and \( T \) is not too high.

Consider a semiconductor with concentrations of acceptors and donors \( N_A \) and \( N_D \). Assume sufficiently high \( T \) that all dopants are ionized. The chemical potential can be found from the **charge neutrality** requirement:

\[
\begin{align*}
n + N_A &= p + N_D \\
np &= n_i^2
\end{align*}
\]

\[
n = \frac{1}{2} \left[ N_D - N_A + \sqrt{(N_D - N_A)^2 + 4n_i^2} \right]
\]

\[
p = \frac{1}{2} \left[ N_A - N_D + \sqrt{(N_A - N_D)^2 + 4n_i^2} \right]
\]

**n-doping** \( n \approx N_D \) if \( N_D \gg n_i, N_A \)

**p-doping** \( p \approx N_A \) if \( N_A \gg n_i, N_D \)

Homework. Check that we still have \( np = n_i^2 \).
**μ in Doped Semiconductors**

Strong $n$-doping pins $\mu$ near the bottom of the conduction band.

$$n = N_D$$

$$n = N_C e^{-\beta(E_C - \mu)}$$

$$\mu = E_C - k_B T \ln \left( \frac{N_D}{N_C} \right)$$

The electron density and the chemical potential of a doped semiconductor as a function of inverse temperature.
At sufficiently low $T$, the electrons/holes recombine with their “parents” (charged donors/acceptors) – **carrier freeze-out**.

How large is $E_C - \mu(T = 0)$ (left panel)?
electron concentration in the intrinsic region

degenerate semiconductor
Problem. Silicon is doped with donors, $N_D = 10^{23}m^{-3}$. This donor concentration is much smaller than $N_D = 10^{25}m^{-3}$ that corresponds to the degenerate case. Assume that all donors are ionized at room temperature. It is also given that $n_i^2 = 10^{32} m^{-6}$.

What are the electron and hole concentrations at (a) 300K, and (b) 0K? Provide short explanations.

(a) 300K

\[
\begin{align*}
n &= \frac{1}{2} \left[ N_D + \sqrt{(N_D)^2 + 4n_i^2} \right] \approx 10^{23}m^{-3} \\
p &= \frac{1}{2} \left[ -N_D + \sqrt{(N_D)^2 + 4n_i^2} \right] \approx \frac{n_i^2}{N_D} = 10^9m^{-3}
\end{align*}
\]

(b) 0K  $n = p = 0$
Conductivity of Doped Semiconductors

\[ \sigma = en\mu \]  
In metals, \( n(T) = const \) and \( \sigma(T) \propto \mu(T) \).

In non-degenerate semiconductors, \( n(T) \propto e^{-\frac{E_C-\mu}{k_BT}} \) dominates in \( \sigma(T) \).

Electron mobility in doped Si

\[ \frac{1}{\mu} = \frac{1}{\mu_{ph}} + \frac{1}{\mu_{imp}} \]
Summary

➢ Intrinsic and doped semiconductors

➢ Doping methods

➢ Law of Mass Action

➢ Chemical potential in semiconductors

➢ Conductivity of intrinsic and doped semiconductors

Today: Simon Ch. 17, Ibach and Luth 12.1-12.5
Next time: MOSFETs