Lecture 18

Motion and Recombination of Electrons and Holes

The motion of carriers in semiconductors is an important topic about semiconductor physic, which dominates properties such as electrical conduction and photoluminescence. The difference between the band structures of CB and VB leads to the treatment of two distinguishable carriers – electron and hole, both contributing to the total electrical conduction. In this lecture, we will introduce several important concepts dealing with the motion and recombination of electrons and holes.

18.1 Drude model of electric current

How do carriers move inside a semiconductor crystal? There are more than one components contributing to the motion of carriers, including the thermal motion, drift (caused by external field) and diffusion (caused by concentration gradient). The thermal motion is present even without external field, and the carriers are moving randomly with a thermal velocity v_{th}:

$$v_{\mathrm{th}} = \sqrt{\frac{3k_{\mathrm{B}}T}{m^*}} \tag{18.1}$$

which combines the form of thermal kinetic energy $E_{\mathrm{k}} = 1.5k_{\mathrm{B}}T = \frac{1}{2}m^*v_{\mathrm{th}}^2$. For silicon at 300 K, $m_{\mathrm{n}}^* = 0.26\ m_e$, which gives $v_{\mathrm{th}} \approx 2.3\times10^5$ m/s. However such motion is non-directional, and an electron easily loses its momentum through collision with the positively-charged lattice (called **scattering**), as shown in Figure 18.1a. Depending on the temperature and the material, the frequency of such events can be characterized by the **mean free time** τ_{m} of the carrier, which is typically at the order of 10^{-13} s for an semiconductor at room temperature. The corresponding **mean free path** is calculated as $l_{\mathrm{m}} = v_{\mathrm{th}}\tau_{\mathrm{m}}$, which is at the order of 10 nm. Between two scattering events, the motion of carrier behave like classical "projectile", and thus often referred as **ballistic transport**. By engineering the quality of material and reducing the temperature, the mean free path inside a semiconductor can be increased, leading to various interesting properties.

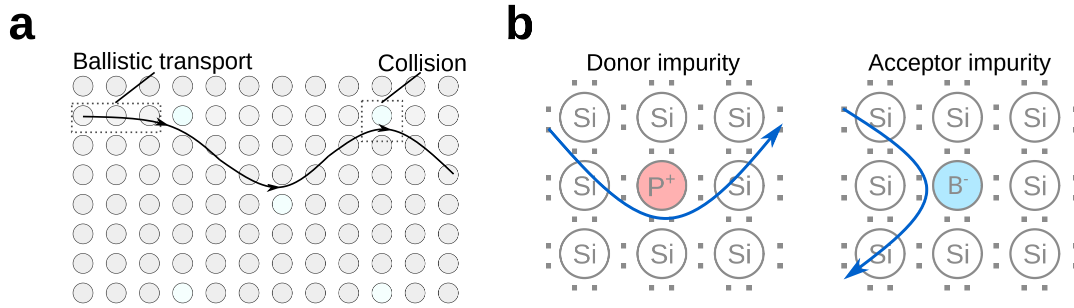

Figure 18.1: Microscopic view of carrier transport inside a semiconductor solid. (a) Ballistic transport and collision (scattering) events at the atomistic level. (b) Scattering of an electron by different charged dopants, showing distinct trajectory.

The *net* thermal motion without external field is always zero. An external electric field accelerates the carriers and give rise to the net drift current, which is calculated by $\mathcal{E}$ is:

$$m^* \frac{\mathrm{d}v_{\mathrm{drift}}}{\mathrm{d}t} = q\mathcal{E} \tag{18.2}$$

where q is the charge of carrier ($-e$ for electron and e for hole). Between the two scattering events, the momentum gained via external force, equals that lost due to the scattering:

$$m^* v_{\mathrm{drift}} = q\mathcal{E}\tau_{\mathrm{m}} \tag{18.3}$$

Similar to the definition of ion mobility we introduced for the electrokinetic phenomena, we can also define the mobility of carriers u, such that $u = |v_{\mathrm{drift}}/q\mathcal{E}|$. Combining with Equation 18.3, the mobilities for electron and hole are defined as:

$$\begin{aligned} u_{\mathrm{n}} &= \frac{e\tau_{\mathrm{m,n}}}{m_{\mathrm{n}}^*} \\ u_{\mathrm{p}} &= \frac{e\tau_{\mathrm{m,p}}}{m_{\mathrm{p}}^*} \end{aligned} \tag{18.4}$$

where $\tau_{\mathrm{m,n}}$ and $\tau_{\mathrm{m,p}}$ are the mean free time for electron and hole, respectively. As seen from the above equations, both electron and hole have mobility proportional to the mean free time, the characteristic time scale of scattering. The scattering of carriers in the crystal is mainly caused by :

1. phonon scattering: collision with the lattice vibration.
2. impurity scattering: elastic scattering with charged, immobile impurities (dopants), changing the trajectory of carrier motion (Figure 18.1b).

By increasing the dopant concentration, the average distance between the dopants reduces, yielding a shorter τ_{m}. For a conducting metal, τ_{m} is typically reduced to the order of fs.[1]

Macroscopically, the drift velocity of carriers leads to the drift current. Consider a 1D semiconductor bar with a voltage drop V across the length L (Figure 18.2). The electrons (blue balls) and holes (red balls) move in opposite directions, and their motion creates current density J_n and J_p (defined as current per unit area), respectively.

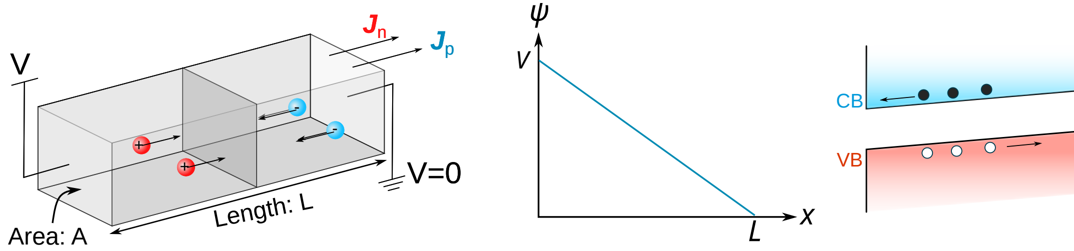

Figure 18.2: Macroscopic view of carrier motion in semiconductor: the Drude model. The potential linearly decreases across the semiconductor which leads to uniform electric field $\mathcal{E}$.

Similar to the mass transfer flux, the **electrical current density** is defined as the amount of charge passes through a unit area per unit time, such that:

$$J_i = z_i e c_i v_\mathrm{drift} \tag{18.5}$$

The total drift current is contributed by both motion of electrons and holes. In 1D system, it is written as:

$$J_\mathrm{drift} = J_\mathrm{n} + J_\mathrm{p} = -nev_\mathrm{drift,n} + pev_\mathrm{drift,p} = (neu_\mathrm{n} + peu_\mathrm{p})\mathcal{E} \tag{18.6}$$

As can be seen, although electrons and holes are moving in opposite directions, the sign of current density is the same. Equation 18.6 indicates that the drift current density is proportional to the electric field. Under the assumption that u_n and u_p are independent of the carrier concentration, it is essentially the Ohm's law:

$$\sigma = \frac{J_\mathrm{drift}}{\mathcal{E}} = neu_\mathrm{n} + peu_\mathrm{p} \tag{18.7}$$

where σ is the electrical conductivity, which has unit of S/m. The inverse of the conductivity is the resistivity $\rho = \sigma^{-1}$ with unit of Ohm·m. The above semi-classical treatment, is usually known as the Drude model of conductivity. Note that Equation 18.7 is only valid in the 1D case. The conductivity and resistivity are in general tensor quantities and does not obey simple inversion, leading to some interesting properties such as the quantum Hall effect (QHE).[2] When the semiconductor is doped, usually the conductivity is governed by the concentration of majority carrier, such that:

- n-doped semiconductor, $\sigma \propto n \sim N_\mathrm{d} - N_\mathrm{a}$
- p-doped semiconductor, $\sigma \propto p \sim N_\mathrm{a} - N_\mathrm{d}$

Equation 18.7 also explains why large bandgap materials are also called insulators. Consider an undoped material with Fermi level lying in the center of the bandgap, the intrinsic carrier density is:

$$n_i = \sqrt{N_c N_c}\exp(-\frac{E_g}{2k_B T}) \tag{18.8}$$

The exponential part decays greatly with E_g. At room temperature, for silicon $E_g = 1.12$ eV, the exponential part is about 5×10^{-10}, still allowing some conductivity. On the other hand, for SiO_2 where $E_g \sim 9$ eV, the exponential part becomes about 10^{-93}, which makes the material insulating.

18.2 Diffusion current

Apart from the drift current which is caused by the external electric field, there is also diffusion current which arises from the gradient of carrier concentration in the semiconductor. The diffusion current density is simply the mass transfer flux multiplied by the carrier charge:

$$J_i = -z_i e\mathcal{D}_i \nabla c_i \tag{18.9}$$

In the 1D case, the diffusion current densities for electron and hole are written as:

$$\begin{aligned} J_{\text{diff,n}} &= e\mathcal{D}_\text{n}\frac{\mathrm{d}n}{\mathrm{d}x} = u_\text{n} k_\text{B} T\frac{\mathrm{d}n}{\mathrm{d}x} \\ J_{\text{diff,p}} &= -e\mathcal{D}_\text{p}\frac{\mathrm{d}p}{\mathrm{d}x} = -u_\text{p} k_\text{B} T\frac{\mathrm{d}p}{\mathrm{d}x} \end{aligned} \tag{18.10}$$

which use the Stokes-Einstein relation, $u_i = \mathcal{D}_i e/k_B T$. The overall current density is the combination of the drift current and the diffusion current, such that:

$$\begin{aligned} J_\text{n} &= J_{\text{drift,n}} + J_{\text{diff,n}} = (k_\text{B} T\frac{\mathrm{d}n}{\mathrm{d}x} - ne\frac{\mathrm{d}\psi}{\mathrm{d}x})u_\text{n} \\ J_\text{p} &= J_{\text{drift,p}} + J_{\text{diff,p}} = (-k_\text{B} T\frac{\mathrm{d}p}{\mathrm{d}x} - pe\frac{\mathrm{d}\psi}{\mathrm{d}x})u_\text{p} \end{aligned} \tag{18.11}$$

where ψ is the electrostatic potential and the relation $\mathcal{E} = -\frac{\mathrm{d}\psi}{\mathrm{d}x}$ is used. Therefore, by knowing the quantities $n(x)$, $p(x)$ and $\psi(x)$, the current densities can be calculated.

18.3 Optical properties of semiconductors

The interaction between light (photon) and carriers is highly related to the band structure, and in particular the bandgap of semiconductors. When the energy of photon $h\nu$ is larger than E_g, the photons are absorbed, exciting electrons at the VB into CB, and leaving holes. Such optical absorption can only occur when $h\nu > E_g$, and ideally the absorption spectrum of a semiconductor is a step function near E_g (Figure 18.3).

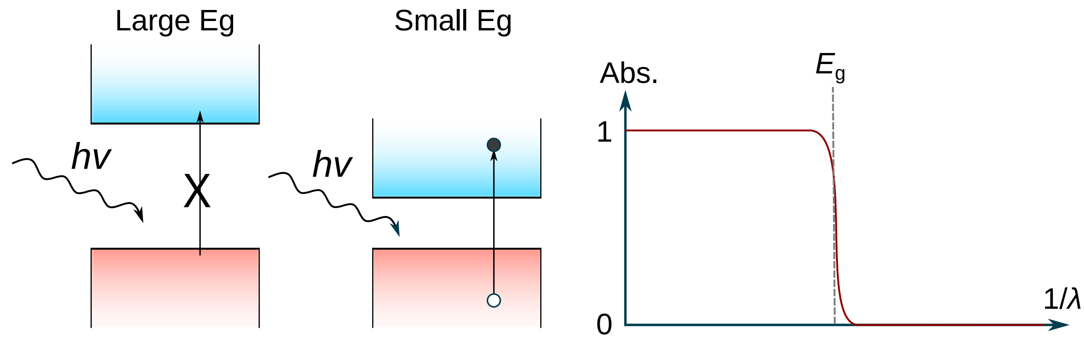

Figure 18.3: Light absorption in a semiconductor. Only the light with wavelength $h\nu$ larger than the bandgap E_g, can be absorbed by the semiconductor. When a photon is absorbed by the semiconductor, a pair of electron-hole is generated. The absorption coefficient α for an ideal semiconductor shows a step near $h\nu \approx E_\mathrm{g}$.

Visible light has wavelength of 380 nm ∼ 780 nm, or equivalently 1.60 eV ∼ 3.26 eV. Therefore, visible light cannot transmit through a metal. The appearance of a semiconductor dependent on its bandgap. For instance, silicon (E_g = 1.12 eV) looks metallic, while GaN (E_g = 3.4 eV) is almost transparent to visible light. For some semiconductors where the minimum of CB and maximum of VB has same momentum (called direct-gap material, such as GaAs, GaN, CdS, etc), the excited electron–hole pair can recombine and generate light, a process known as the photoluminescence (PL), as shown in Figure 18.4a. Even when $h\nu$ is larger than E_g, the excited carriers need to first thermally relax to the edge of CB and VB before recombination, making the energy of the PL peak around E_g, as shown in Figure 18.4b. The efficiency of photon–carrier conversion, is known as the quantum yield (QY), which is an important quantity in photodetectors and solar cells that we will discuss in Lecture 21. An indirect-gap material, like silicon, usually has low QY due to the heat loss. Another important characteristic is the full width at half maximum (FWHM) of the PL peak, a narrower FWHM indicates better semiconductor quality.

The photon–carrier interaction involves the generation and recombination processes. At thermal equilibrium, the electron and hole concentration follows $n \times p = n_i^2$ = Const.. However, when an semiconductor is illustrated by light, additional carriers are generated, such that:

$$\begin{aligned} n &= \underbrace{n_0}_{\text{thermal}} + \underbrace{n'}_{\text{illumination}} \\ p &= \underbrace{p_0}_{\text{thermal}} + \underbrace{p'}_{\text{illumination}} \end{aligned} \tag{18.12}$$

where $n' = p'$. As a result, after illumination, $n * p$ is larger than n_i^2, and thermal equilibrium no longer holds. The generation process: null → e^- + h^+ is zeroth-order and associated with rate g (unit: $cm^{-3}\cdot s^{-1}$). On the other hand, the recombination process e^- + h^+ → null, is a first-order process with a lifetime of τ. The time-dependent continuity equation of carrier

concentration is contributed by flux, generation and recombination:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = -\nabla \cdot J_i + g - \frac{c_i'}{\tau} \tag{18.13}$$

where c_i is n or p. In a 1D system, Equation 18.13 is written as:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = u_i k_\mathrm{B} T \frac{\mathrm{d}^2 c_i}{\mathrm{d}x^2} + u_\mathrm{i} z_i e \frac{\mathrm{d}}{\mathrm{d}x}(c_i \frac{\mathrm{d}\psi}{\mathrm{d}x}) + g - \frac{c_i'}{\tau} \tag{18.14}$$

Without drift and diffusion current ($\mathrm{d}c_i/\mathrm{d}x = 0$ and $\mathrm{d}\psi/\mathrm{d}x = 0$), we know that at steady state, the concentration of n' and p', is related with the generation rate the recombination lifetime:

$$n' = p' = g\tau \tag{18.15}$$

When a semiconductor is illuminated, the concept of Fermi level is no longer valid, since there is no single energy level at which the occupation numbers of electron and hole are both 1/2. In practice, the idea of "quasi" Fermi levels for electron and hole, are introduced, such that:

$$\begin{aligned} n = n_0 + n' = N_\mathrm{c} \exp(-\frac{E_\mathrm{c} - E_\mathrm{Fn}}{k_\mathrm{B} T}) \\ p = p_0 + p' = N_\mathrm{v} \exp(-\frac{E_\mathrm{Fp} - E_\mathrm{v}}{k_\mathrm{B} T}) \end{aligned} \tag{18.16}$$

As can be seen, E_Fn and E_Fp do not coincide (Figure 18.4b), with E_Fn closer to the CB and E_Fp closer to the VB. The concept of quasi Fermi level is important for semiconductor interfaces such as the p-n junction, which will be discussed later.

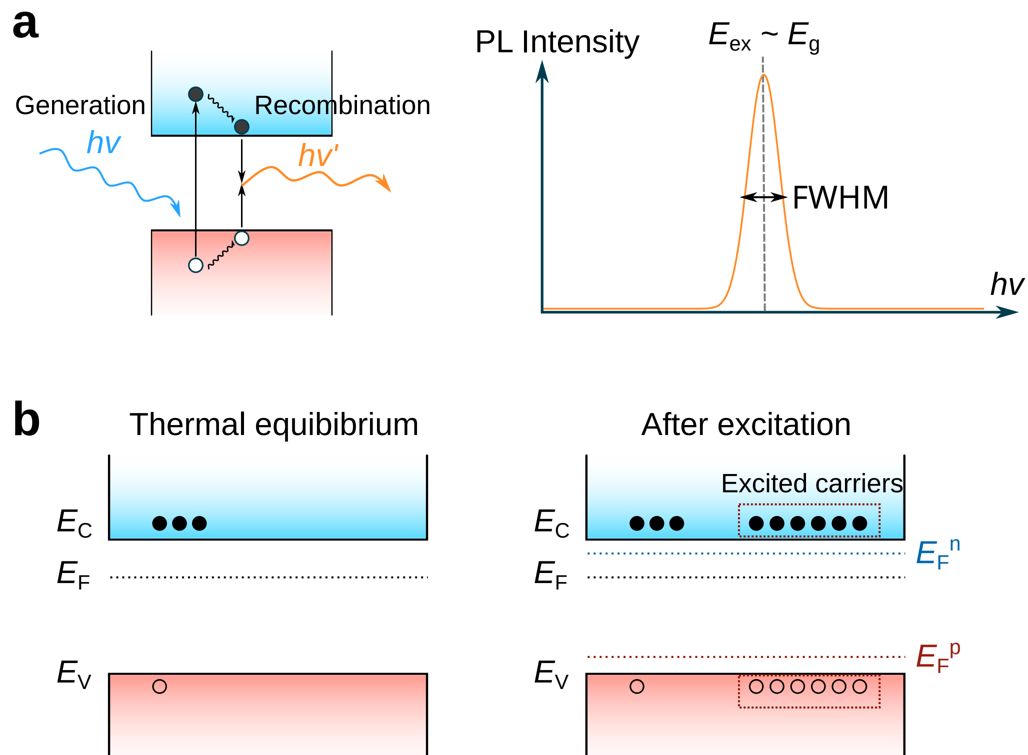

Figure 18.4: Photon-carrier interaction in a semiconductor. (a) Generation and recombination of carriers. During the recombination process, photoluminescence (PL) may be observed according to the band structure of the semiconductor, showing a PL peak around E_g. The full width at half maximum (FWHM) of the PL peak usually characterizes the quality of the semiconductor. (b) Carrier population at thermal equilibrium (left) and after generation (right) for an n-doped semiconductor. The excess carriers n' and p' cause the emergence of "quasi" Fermi levels E_{Fn} and E_{Fp}.

References

(1) Gall, D. *J. Appl. Phys.* **2016**, *119*, 085101.

(2) Cage, M. E.; Klitzing, K.; Chang, A.; Duncan, F.; Haldane, M.; Laughlin, R.; Pruisken, A.; Thouless, D., *The quantum Hall effect*; Springer Science & Business Media: 2012.