Solid-State Electronics
EE230C

Class Notes for EE230C taught by Prof Sayeef Salahuddin
Scribed by Yen Kai Lin in 2018
Further Edited and Augmented by Pratik Brahma in 2023

September 7, 2023

Contents

Chapter 1

General Considerations of Electronic Transport

1.1 Ohm's Law

Ohm's law describes the relationship between the current (I) and applied voltage (V) or electric field (E) in a conductor and it can be expressed as

$$I = \frac{V}{R} \tag{1.1}$$

where R is the resistance and is a function of the size of conductor

$$R = \frac{\rho L}{A} = \frac{L}{A}\frac{1}{\sigma} \tag{1.2}$$

Hence, Ohm's law can be rewritten as

$$I = \frac{V\sigma A}{L} \Rightarrow J = \frac{I}{A} = V\frac{\sigma}{L} = \sigma E \tag{1.3}$$

σ is the conductivity and it's related to the mobility ($\mu = \frac{e\tau}{m^*}$).

$$\sigma = ne\mu = \frac{e^2 n\tau}{m^*} \tag{1.4}$$

From Equation (1.1) to (1.4), one may find that if L is very small or τ (scattering or collision time) approaches to ∞ (happens when the temperature is close to zero), then the conductivity or current (density) would become huge without any bound. Therefore, the question is: What does the current look like in a very small device? The experimental result[1] has shown that the conductance of a point contact is quantized at a low temperature, which seems to contradict Ohm's law shown above. Landauer's formula will be discussed in the next section to explore this issue.

---

[1] Quantized Conductance of Point Contacts in a Two-Dimensional Electron Gas, B. J. van Wees et al., PRL, 1988
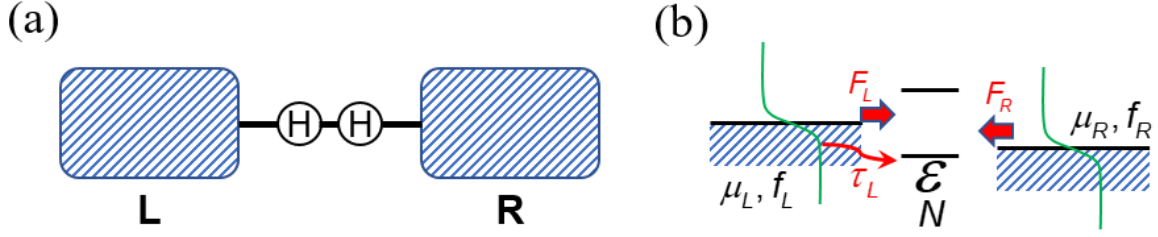
Figure 1.1: (a) Schematic diagram of the "hydrogen molecule" device with left (L) and right (R) contact. (b) Energy diagram when applying a bias $-eV = \mu_R - \mu_L$. $F_{L,R}$ and $f_{L,R}$ are the electron flux and Fermi function of the left and right contacts. The green curves indicate the probability distribution of finding an electron (Fermi distribution).

1.2   Landauer's Formula

Consider a device consisting of a hydrogen molecule (with two energy levels) connected between two contacts, as shown in Fig. 1.1. We assume that when applying a bias V, the electrons can flow through the hydrogen molecule without scattering. Also, the Fermi function (distribution) is not applicable for the hydrogen molecule because the amount of electrons is not sufficient to be statistically meaningful[2]. However, the contacts (like electron reservoirs) are assumed to be "large" enough to support the Fermi function. The equation of states (continuity equation) at a given energy ε_i can be written as

$$\frac{dN^{\varepsilon_i}}{dt} = F_L^{\varepsilon_i} + F_R^{\varepsilon_i} \tag{1.5}$$

where N^{ϵ_i} is the occupation factor at energy ϵ_i and the fluxes are

$$F_L^{\varepsilon_i} = \frac{f_L^{\varepsilon_i} - N^{\varepsilon_i}}{\tau_L^{\varepsilon_i}} \tag{1.6}$$

$$F_R^{\varepsilon_i} = \frac{f_R^{\varepsilon_i} - N^{\varepsilon_i}}{\tau_R^{\varepsilon_i}} \tag{1.7}$$

Hence, Equation (1.5) becomes[3]

$$\frac{dN}{dt} + N\left(\frac{1}{\tau_L} + \frac{1}{\tau_R}\right) = \frac{f_L}{\tau_L} + \frac{f_R}{\tau_R} \tag{1.8}$$

Now, we define[4]

$$\frac{1}{\tau_{L,R}} = \frac{\gamma_{L,R}}{\hbar} \tag{1.9}$$

---

[2]Ashley H. Carter, Classical and Statistical Thermodynamics.
[3]The superscript ε_i is omitted for simplicity.
[4]The physical meaning will be discussed later.

where γ has the unit of energy. Thus, we get

$$\hbar \frac{dN}{dt} + N\left(\gamma_L + \gamma_R\right) = \gamma_L f_L + \gamma_R f_R \tag{1.10}$$

At steady state, we have

$$N = \frac{\gamma_L f_L + \gamma_R f_R}{\gamma_L + \gamma_R} \tag{1.11}$$

Therefore, the current at energy ε_i can be expressed as

$$
\begin{aligned}
I^{\varepsilon_i} = e F_L^{\varepsilon_i} &= e \frac{f_L^{\varepsilon_i} - N^{\varepsilon_i}}{\tau_L^{\varepsilon_i}} \\
&= \frac{e}{\hbar} \gamma_L^{\varepsilon_i} \left(f_L^{\varepsilon_i} - \frac{\gamma_L^{\varepsilon_i} f_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i} f_R^{\varepsilon_i}}{\gamma_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i}} \right) \\
&= \frac{e}{\hbar} \frac{\gamma_L^{\varepsilon_i} \gamma_R^{\varepsilon_i}}{\gamma_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i}} \left(f_L^{\varepsilon_i} - f_R^{\varepsilon_i} \right)
\end{aligned}
\tag{1.12}
$$

From Equation (1.12), one may directly observe that the current only depends on the properties of the contacts. Now if there are multiple energy levels, the current becomes

$$I = \sum_{\varepsilon_i} I^{\varepsilon_i} \times 2 = \sum_{\varepsilon_i} \frac{e}{\hbar} \frac{\gamma_L^{\varepsilon_i} \gamma_R^{\varepsilon_i}}{\gamma_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i}} \left(f_L^{\varepsilon_i} - f_R^{\varepsilon_i} \right) \times 2 \tag{1.13}$$

where 2 is coming from the spin degeneracy. Use the following trick

$$g\left(\varepsilon_i\right) = \int dE\, g\left(E\right) \delta\left(E - \varepsilon_i\right) \tag{1.14}$$

and Equation (1.13) can be rewritten as

$$
\begin{aligned}
I &= \int dE \left(\frac{2e}{\hbar}\right) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \left(f_L^E - f_R^E \right) \sum_{\varepsilon_i} \delta\left(E - \varepsilon_i\right) \\
&= \frac{2e}{\hbar} \int dE\, D\left(E\right) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \left(f_L^E - f_R^E \right)
\end{aligned}
\tag{1.15}
$$

where Equation (1.15) is **Landauer's formula** and $D\left(E\right)$ is the density of states (DOS).

$$D\left(E\right) = \sum_{\varepsilon_i} \delta\left(E - \varepsilon_i\right) \tag{1.16}$$

The general definition of DOS is the number of states per unit energy, so $D\left(E\right)$ also can be written as

$$D\left(E\right) = \frac{\partial N_E^{CS}}{\partial E} = \lim_{\Delta E \to 0} \frac{N_E^{CS}\left(E + \Delta E\right) - N_E^{CS}\left(E\right)}{\Delta E} \tag{1.17}$$

where N_E^{CS} is[5]

$$N_E^{CS} = \sum_{\varepsilon_i} \theta\left(E - \varepsilon_i\right) \tag{1.18}$$
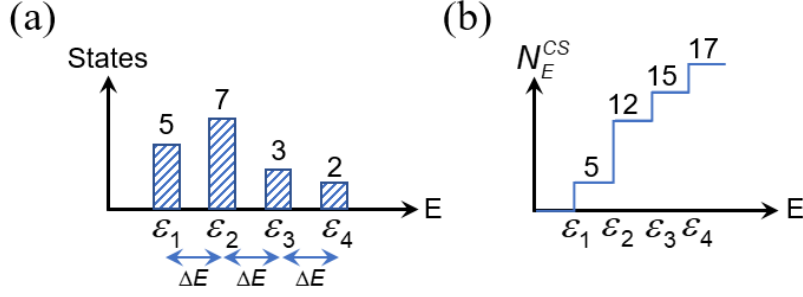
Figure 1.2: (a) Number of states distribution. (b) Cumulative sum of states.

where $\theta\left(E-\varepsilon_i\right)$ is the Heaviside step function with the implicit degeneracy as shown in Fig. 1.2. For example, at energy ε_1 there are 5 states [Fig. 1.2(a)]. As it goes from low energy to high energy, the cumulative sum of states increases from 5 to 17 [Fig. 1.2(b)]. Thus, the Equation (1.17) becomes

$$D\left(E\right) = \frac{\partial N^{CS}}{\partial E} = \sum_{\varepsilon_i} \delta\left(E-\varepsilon_i\right) \tag{1.19}$$

where $\delta\left(E-\varepsilon_i\right)$ is the delta function. After knowing the DOS, we obtain the total occupation factor N

$$
\begin{aligned}
N &= \sum_{\varepsilon_i} \frac{\gamma_L^{\varepsilon_i} f_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i} f_R^{\varepsilon_i}}{\gamma_L^{\varepsilon_i} + \gamma_R^{\varepsilon_i}} \\
&= \int dE \frac{\gamma_L^E f_L^E + \gamma_R^E f_R^E}{\gamma_L^E + \gamma_R^E} \sum_{\varepsilon_i} \delta\left(E-\varepsilon_i\right) \\
&= \int dE \frac{\gamma_L^E f_L^E + \gamma_R^E f_R^E}{\gamma_L^E + \gamma_R^E} D\left(E\right)
\end{aligned}
\tag{1.20}
$$

Equation (1.20) is under the non-equilibrium condition. At the equilibrium condition (applied voltage $V = 0$), the Fermi functions and γ of both contacts are the same ($f_L^E = f_R^E = f_0^E, \gamma_L^E = \gamma_R^E, \forall E$). Thus, the Equation (1.20) can be re-written as

$$N = \int_{-\infty}^{\infty} D\left(E\right) f_0\left(E\right) dE \tag{1.21}$$

Equation (1.21) is commonly used to calculate the carrier concentration in semiconductors. Now, we would like to figure out the maximum conductance based on Landauer's formula. Two assumptions are made in the following discussion: (1) the applied voltage V is very small and (2) the temperature is fairly low. At small V, the Taylor expansion with the first two terms can be applied to the Fermi functions.

$$f_L = f_0\left(E-\mu_L\right) \tag{1.22}$$

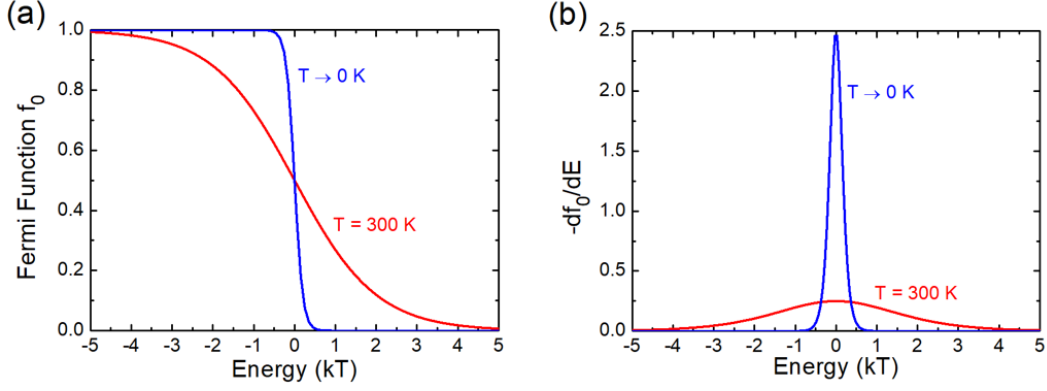[5]Here is the cumulative sum (CS) of states at and below a given energy E.

Figure 1.3: (a) Fermi functions and (b) their derivative at different temperatures. The Fermi level is located at $E = 0$.

$$f_R = f_0 (E - \mu_R)$$
$$= f_0 (E + eV - \mu_L)$$
$$= f_0 (E - \mu_L) + \left(\frac{\partial f_0}{\partial E} \right) eV \tag{1.23}$$

Therefore, the Landauer's formula becomes

$$I = \frac{2e}{\hbar} \int dE D(E) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \left(f_L^E - f_R^E \right)$$
$$\approx \frac{2e^2}{\hbar} V \int dE D(E) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \left(\frac{-\partial f_0}{\partial E} \right) \tag{1.24}$$

Hence, the conductance G is

$$G = \frac{\partial I}{\partial V} = \frac{2e^2}{\hbar} \int dE D(E) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \left(\frac{-\partial f_0}{\partial E} \right) \tag{1.25}$$

When the temperature approaches zero, the derivative of the Fermi function with respect to the energy E is like a delta function (see Fig. 1.3)

$$\left(\frac{-\partial f_0 (E - \mu_L)}{\partial E} \right) \rightarrow \delta (E - \mu_L) \tag{1.26}$$

and the conductance G becomes

$$G = \frac{2e^2}{\hbar} \int dE D(E) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \delta (E - \mu_L)$$
$$= \frac{2e^2}{\hbar} \left[D(E) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \right]_{E = \mu_L} \tag{1.27}$$

To further examine the Equation (1.27), the concept of "energy broadening" should be introduced.

1.3   Broadening and Maximum Conductance

Due to the wave nature of the quantum objects, the uncertainty principle is inherent in the system we discussed. The uncertainty principle states that complementary variables, such as position/momentum and energy/time, cannot be measured simultaneously[6].

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{1.28}$$

The Δt can be interpreted as the "lifetime" of electrons, implying that the ΔE is non-zero, which is called energy "broadening." This effect has been seen in Equation (1.9) where γ is interpreted as the energy broadening. When the atomistic device interacts with a large quantum system (L and R contacts), the energy levels no longer remain discrete but change to a spectrum-like distribution (broadening). However, the shape of the energy broadening is difficult to be experimentally determined, so we utilize the Lorentzian equation[7] to get analytic results

$$\delta \left(E - \varepsilon_i \right) \xrightarrow{\text{Broadening}} \frac{1}{2\pi} \frac{\gamma_L + \gamma_R}{\left(E - \varepsilon_i \right)^2 + \left(\frac{\gamma_L + \gamma_R}{2} \right)^2} \tag{1.29}$$

For the Lorentzian equation, $\gamma_{L/R}$ are constants and not a function of energy. Then Equation (1.29) gives

$$\int_{-\infty}^{\infty} dE \frac{1}{2\pi} \frac{\gamma_L + \gamma_R}{\left(E - \varepsilon_i \right)^2 + \left(\frac{\gamma_L + \gamma_R}{2} \right)^2} = \frac{\gamma_L + \gamma_R}{2\pi} \frac{2}{\gamma_L + \gamma_R} \arctan \left(\frac{E - \varepsilon_i}{\frac{\gamma_L + \gamma_R}{2}} \right) \Bigg|_{-\infty}^{\infty} = 1 \tag{1.30}$$

Therefore, the Equation (1.27) can be re-written as

$$
\begin{aligned}
G &= \frac{2e^2}{\hbar} \sum_{\varepsilon_i} \delta \left(E - \varepsilon_i \right) \frac{\gamma_L^E \gamma_R^E}{\gamma_L^E + \gamma_R^E} \Bigg|_{E = \mu_L} \\
&\xrightarrow{\text{Broadening}} \frac{2e^2}{\hbar} \sum_{\varepsilon_i} \frac{1}{2\pi} \frac{\gamma_L + \gamma_R}{\left(\mu_L - \varepsilon_i \right)^2 + \left(\frac{\gamma_L + \gamma_R}{2} \right)^2} \frac{\gamma_L \gamma_R}{\gamma_L + \gamma_R} \\
&= \frac{2e^2}{h} \sum_{\varepsilon_i} \frac{4 \gamma_L \gamma_R}{4 \left(\mu_L - \varepsilon_i \right)^2 + \left(\gamma_L + \gamma_R \right)^2} \\
&\xrightarrow[\mu_L = \varepsilon_i]{maximum} \frac{2e^2}{h} \frac{4 \gamma_L \gamma_R}{\left(\gamma_L + \gamma_R \right)^2} M
\end{aligned} \tag{1.31}
$$

M is the number of degenerate states at energy ε_i called the number of modes. From the inequality of arithmetic and geometric means[8], the maximum conductance is

$$G^{max} = \frac{2e^2}{h} M \tag{1.32}$$

---

[6]D. J. Griffiths, Introduction to Quantum Mechanics, Chapter 3.

[7]We would like a function with (1) symmetry and (2) area under the curve $= 1$.

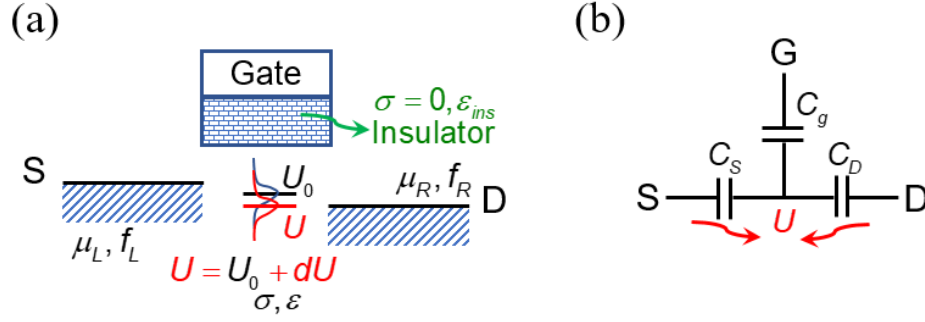[8]$\frac{a+b}{2} \geq \sqrt{ab} \Rightarrow \frac{4ab}{(a+b)^2} \leq 1$

Figure 1.4: (a) Schematic diagram of the gated device. The energy level in the channel is broadened due to the current flow. (b) The equivalent circuit of the gated device. U and U_0 are the energy with and without biases.

Equation (1.32) indicates that the material band structure and dimensionality limit the maximum conductance. This is also why plateaus are observed in the conductance measurement at low temperatures.

1.4 Electrostatics

In the previous sections, the device with two terminals is discussed. Now, let's consider a three-terminal device (like MOSFET) with an insulator layer and a gate contact on the top of the channel as shown in Fig. 1.4. When the gate voltage is applied ($V_g \neq 0, V_S = V_D = 0$), the energy level in the channel is pulled down due to more electrons induced and the change in the electrostatic potential (ϕ) in the channel can be expressed as

$$d\phi = \frac{C_g}{C_g + C_S + C_D} dV_g \tag{1.33}$$

and the occupation factor with (N) and without bias (N_0) are

$$N = \int_{-\infty}^{\infty} dE \frac{\gamma_L f_L + \gamma_R f_R}{\gamma_L + \gamma_R} D(E - U) \tag{1.34}$$

$$N_0 = \int_{-\infty}^{\infty} dE \frac{\gamma_L f_L + \gamma_R f_R}{\gamma_L + \gamma_R} D(E - U_0) \tag{1.35}$$

$$\delta N \equiv N - N_0 \tag{1.36}$$

Similarly, if $V_S \neq 0, V_g = V_D = 0$, the potential change is

$$d\phi = \frac{C_S}{C_g + C_S + C_D} dV_S \tag{1.37}$$

and if $V_D \neq 0, V_g = V_S = 0$,

$$d\phi = \frac{C_D}{C_g + C_S + C_D} dV_D \tag{1.38}$$

Finally, if $V_g = V_S = V_D = 0$, and Q is the amount of charge present in the channel,

$$d\phi = \frac{dQ}{C_g + C_S + C_D} \tag{1.39}$$

Combining above equations, we obtain

$$d\phi = \alpha_g dV_g + \alpha_D dV_D + \alpha_S dV_S + \frac{dQ}{C_g + C_S + C_D} \tag{1.40}$$

where

$$\alpha_{g,S,D} = \frac{C_{g,S,D}}{C_g + C_S + C_D} \tag{1.41}$$

The Equation (1.40) also can be expressed using energy U (Multiplying $-e$ to the above equation)

$$dU = \alpha_g U_g + \alpha_D U_D + \alpha_S U_S - \frac{edQ}{C_g + C_S + C_D}$$

$$= \alpha_g U_g + \alpha_D U_D + \alpha_S U_S + \frac{e^2 \delta N}{C_g + C_S + C_D} \tag{1.42}$$

The Equation (1.42) tells us that when increasing the number of electrons (occupation) in the channel, there is an energy penalty that resists the changes. Solving Equation (1.34), (1.35), (1.36), and (1.42) by iterations, the energy U is obtained and the current can be immediately calculated using Landauer's formula

$$I = \frac{2e}{\hbar} \int dE D (E - U) \frac{\gamma_L \gamma_R}{\gamma_L + \gamma_R} (f_L - f_R) \tag{1.43}$$

where we define

$$D (E - U) \frac{\gamma_L \gamma_R}{\gamma_L + \gamma_R} \equiv T (E) \tag{1.44}$$

is the transmission probability due to scatterings. Assuming the applied drain-to-source voltage (V) is small and the energy broadening follows Equation (1.29), we have

$$I = \frac{2e}{\hbar} \sum_{\varepsilon_i} \int dE \frac{1}{2\pi} \frac{\gamma_L + \gamma_R}{(E - \varepsilon_i)^2 + \left(\frac{\gamma_L + \gamma_R}{2}\right)^2} \frac{\gamma_L \gamma_R}{\gamma_L + \gamma_R} \left(\frac{-\partial f_0}{\partial E}\right) eV$$

$$= \frac{2e^2}{h} V \sum_{\varepsilon_i} \underbrace{\int dE \times \frac{\gamma_L \gamma_R}{(E - \varepsilon_i)^2 + \left(\frac{\gamma_L + \gamma_R}{2}\right)^2} \left(\frac{-\partial f_0}{\partial E}\right)}_{\equiv \overline{T}_{\varepsilon_i}}$$

$$= \frac{2e^2}{h} V \overline{M} \cdot \overline{T}_{\varepsilon_i} \tag{1.45}$$

where $\overline{M}$ is the average number of modes around the Fermi energy. The resistance is

$$R = \frac{\partial V}{\partial I} = \frac{h}{2e^2 \overline{M}} \cdot \frac{1}{\overline{T}_{\varepsilon_i}} \tag{1.46}$$

Chapter 2

Density of States

2.1 Schrodinger Equation

Although the classical physics was well developed in 1900s, the black body radiation implicitly indicated that something is missing in the classical physics. From classical physics, the more heat putting in an object would increase the kinetic energy of the object, and thus the energy of the emitted radiation should increase with the frequency which is known as "ultraviolet catastrophe." However, this contradicted with the experimental results that the radiation intensity tends to be zero at short wavelength (higher frequency) regime. To explain this black body radiation, the (implicit) emitted energy quantization and concept of photon (particle of light) were proposed by Planck in 1900[1]. After that, physicists gradually noticed that every object exhibits both wave-particle duality and the energy states of a system are quantized. Finally, in 1925 Schrodinger wrote down an equation analogous to a wave equation: Schrodinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m_0} \frac{\partial^2}{\partial x^2} \Psi + V\Psi \tag{2.1}$$

where Ψ is the wavefunction of a particle, m_0 is the rest mass of the particle, and V is the potential energy experienced by a particle.

2.1.1 Time-Independent Schrodinger Equation

The time independent Schrodinger equation[2] predicts that the wavefunctions are able to form the standing waves which is called stationary states (also called "orbitals," such as atomic orbitals and molecular orbitals). In stationary states, the wavefunction can be divided into two parts using separation of variation method:

$$\Psi = \psi(x)\varphi(t) \tag{2.2}$$

---

[1] Einstein explicitly assumed that the electromagnetic radiation is quantized in 1905.
[2] D. J. Griffiths, Introduction to Quantum Mechanics, Chapter 2.

where ψ is a function of x alone, and φ is a function of t alone. There, we can rewrite Equation (2.1)

$$i\hbar\psi\frac{\partial\varphi}{\partial t} = -\frac{\hbar^2}{2m_0}\frac{\partial^2\psi}{\partial x^2}\varphi + V\psi\varphi \tag{2.3}$$

Dividing through by $\psi\varphi$:

$$i\hbar\frac{1}{\varphi}\frac{\partial\varphi}{\partial t} = -\frac{\hbar^2}{2m_0}\frac{1}{\psi}\frac{\partial^2\psi}{\partial x^2} + V \tag{2.4}$$

Now, the left side and the right side are functions of t and x, respectively[3]. Thus, both sides are equqal to a constant E. Then

$$i\hbar\frac{1}{\varphi}\frac{\partial\varphi}{\partial t} = E \tag{2.5}$$

and

$$-\frac{\hbar^2}{2m_0}\frac{1}{\psi}\frac{\partial^2\psi}{\partial x^2} + V = E \tag{2.6}$$

The general solution of Equation (2.5) is

$$\varphi(t) = e^{-iEt/\hbar} \tag{2.7}$$

and Equation (2.6) becomes

$$-\frac{\hbar^2}{2m_0}\frac{\partial^2\psi}{\partial x^2} + V\psi = E\psi$$

Or

$$\boxed{\left[-\frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial x^2} + V\right]\psi = E\psi} \tag{2.8}$$

The Equation (2.8) is called the **time-independent Schrodinger equation** and

$$-\frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial x^2} + V \equiv H \tag{2.9}$$

is called the **Hamiltonian** which is an operator. The Equation (2.8) implicitly states that solving the Schrodinger equation is an eigenvalue problem. The physical meaning of the constant E is actually the energy of the particle. Thus, the energy available in a quantum system can be determined by Equation (2.8). Taking a simplest example of a free electron ($V = 0$), we can immediately solve Schrodinger equation and get

$$\psi = Ae^{ikx} \tag{2.10}$$

and

$$E = \frac{\hbar^2 k^2}{2m_0} \tag{2.11}$$

---

[3]Here, V is only a function of x.

where k is the wavenumber. Note that the Equation (2.10) is a plane wave solution. If we apply the periodic boundary condition, we have

$$\psi\left(x = 0\right) = \psi\left(x = L\right)$$

$$1 = e^{ikL} \Rightarrow kL = 2\pi n \Rightarrow k = \frac{2\pi}{L}n \tag{2.12}$$

and

$$E = \frac{\hbar^2}{2m_0} \frac{4\pi^2}{L^2} n^2 \tag{2.13}$$

where n is an integer. Thus, for an one-dimensional free electron, the density of states is

$$
\begin{aligned}
D\left(E\right) &= \sum_n \delta\left(E - E_n\right) = \sum_n \delta\left(E - \frac{\hbar^2}{2m_0}\frac{4\pi^2}{L^2}n^2\right) \\
&= \int_{-\infty}^{\infty} \frac{dk}{\Delta k} \delta\left(E - \frac{\hbar^2 k^2}{2m_0}\right), \quad \frac{\hbar^2 k^2}{2m_0} \equiv z, \quad \Delta k = \frac{2\pi}{L} \\
&= \frac{Lm_0}{\pi\hbar}\frac{1}{\sqrt{2m_0}}\int_0^{\infty} dz \frac{1}{\sqrt{z}}\delta\left(E - z\right) \\
&= \frac{Lm_0}{\pi\hbar}\frac{1}{\sqrt{2m_0}}\frac{1}{\sqrt{E}}
\end{aligned} \tag{2.14}
$$

Note that the electron has up- and down-spin degeneracy, so the total DOS (per unit length per unit energy) is

$$\boxed{\frac{D\left(E\right)}{L} \equiv D\left(E\right) = \frac{2m_0}{\pi\hbar}\frac{1}{\sqrt{2m_0}}\frac{1}{\sqrt{E}}} \tag{2.15}$$

2.1.2 Formalism in Quantum Mechanics

The formalism in quantum mechanics is somehow different from the classical physics. There are two important constructs: wavefunctions and operators, which represent the state of a system and observables. In mathematical language (linear algebra), the wavefunctions are vectors and, the operators are the linear transformations. A vector in an N-dimensional space can be expressed using a specified orthogonal basis

$$\left|\psi\right\rangle \rightarrow \boldsymbol{\psi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} \tag{2.16}$$

and

$$\left\langle\psi\right| \rightarrow \boldsymbol{\psi}^\dagger = \begin{bmatrix} \phi_1^* & \phi_2^* & \cdots & \phi_N^* \end{bmatrix} \tag{2.17}$$

The Equations (2.16) and (2.17) are the ket and bra, which form the **Dirac notation**. In general, the solution of Schrodinger equation is

$$|\psi\rangle = \sum_{j}^{N} c_j |\phi_j\rangle \tag{2.18}$$

and ϕ_j's form a set of orthogonal basis in N-dimensional **Hilbert space** obeying

$$\int_{-\infty}^{\infty} \phi_i^* (x) \, \phi_j (x) \, dx = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{2.19}$$

where ϕ_i^* is the complex conjugate of ϕ_i and the Equation (2.19) is the inner product of two bases. The coefficients c_j's can be obtained by

$$\boxed{\langle \phi_k | \psi \rangle = \int_{-\infty}^{\infty} \phi_k^* [c_1 \phi_1 + c_2 \phi_2 + \ldots + c_k \phi_k + \ldots] \, dx = c_k} \tag{2.20}$$

Using Dirac notation to rewrite Schrodinger equation, we have[4]

$$E|\psi\rangle = \boldsymbol{H}|\psi\rangle \tag{2.21}$$

If multiplying $\langle \phi_k |$ to the Equation (2.21),

$$\sum_{j} E c_j \langle \phi_k | \phi_j \rangle = \sum_{j} \langle \phi_k | \boldsymbol{H} c_j | \phi_j \rangle$$

$$\Rightarrow E c_k = \sum_{j} \langle \phi_k | \boldsymbol{H} | \phi_j \rangle c_j = \sum_{j} H_{k,j} c_j \tag{2.22}$$

where

$$H_{k,j} \equiv \langle \phi_k | \boldsymbol{H} | \phi_j \rangle \tag{2.23}$$

Thus,

$$E \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} H_{11} c_1 + H_{12} c_2 + \ldots \\ H_{21} c_1 + H_{22} c_2 + \ldots \\ \vdots \end{bmatrix} = \boldsymbol{H}\,[c]$$

$$\Rightarrow \boxed{E\,[c] = \boldsymbol{H}\,[c]} \tag{2.24}$$

The Equation (2.24) explicitly indicates that solving the Schrodinger equation is mathematically an eigenvalue problem, where $[c]$ is the eigenvectors with the corresponding eigenvalues E.

---

[4]The bold font represents an operator here and later on. For example, the Hamiltonian is an operator and the energy E is a scalar.

2.1.3 Basis Transformation

As seen in Equation (2.18), there are multiple choices of the set of basis. The question is: how can we change the basis and what are the mathematical and physical meanings of such transformation? Let's consider

$$|\psi\rangle = \sum_j c_j |\phi_j\rangle = \sum_k d_k |u_k\rangle \tag{2.25}$$

$$\Rightarrow d_i = \langle u_i|\psi\rangle = \sum_j \langle u_i|\phi_j\rangle c_j \equiv \sum_j M_{ij} c_j \tag{2.26}$$

$$\Rightarrow \begin{bmatrix} d_1 \\ d_2 \\ \vdots \end{bmatrix} = [M] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} \tag{2.27}$$

where M is the rotation matrix. Mathematically, the Equation (2.27) represents the rotation of a vector:

$$A|x\rangle = |y\rangle \tag{2.28}$$

and

$$\langle y| = \langle x|A^\dagger \tag{2.29}$$

such taht

$$\langle y|A|x\rangle = \langle y|y\rangle$$
$$\Rightarrow \langle x|A^\dagger A|x\rangle = \langle y|y\rangle = \langle x|x\rangle = 1 \tag{2.30}$$

Therefore,

$$A^\dagger A = I \tag{2.31}$$

where I is the identity matrix and the Equation (2.31) is called unitary transformation. Now if we have a two-dimensional vector

$$|r\rangle = a|x\rangle + b|y\rangle \tag{2.32}$$

where

$$a = \langle x|r\rangle \quad \text{and} \quad b = \langle y|r\rangle$$

then

$$|r\rangle = |x\rangle a + |y\rangle b = |x\rangle\langle x|r\rangle + |y\rangle\langle y|r\rangle$$
$$= \underbrace{(|x\rangle\langle x| + |y\rangle\langle y|)}_{=1} |r\rangle \tag{2.33}$$

More generally,

$$\sum_i |\phi_i\rangle\langle\phi_i| = 1 \tag{2.34}$$

The Equation (2.34) is called the identity operator. By using the identity operator, we obtain the new Hamiltonian in a new set of basis.

$$
\begin{aligned}
H_{kj} &= \langle u_k | \boldsymbol{H} | u_l \rangle \\
&= \sum_{i,j} \langle u_k | \phi_i \rangle \langle \phi_i | \boldsymbol{H} | \phi_j \rangle \langle \phi_j | u_l \rangle \\
&= \sum_{i,j} M_{ki} H_{ij} M_{jl}
\end{aligned}
$$

$$
\Rightarrow \boldsymbol{M}^\dagger \boldsymbol{H}^{\text{old}} \boldsymbol{M} = \boldsymbol{H}^{\text{new}} \tag{2.35}
$$

As mentioned earlier, solving the Schrodinger equation is a eigenvalue problem. Thus, the property of the Hamiltonian should be properly explored. Suppose we have the following eigenvalue problem

$$
\boldsymbol{A} | x \rangle = \lambda | x \rangle \quad \text{and} \quad \langle x | \boldsymbol{A}^\dagger = \lambda^\dagger \langle x | \tag{2.36}
$$

By multiplying $\langle x |$ and $| x \rangle$ from the left and right sides respectively to the Equation (2.36), we have

$$
\langle x | \boldsymbol{A} | x \rangle = \lambda \quad \text{and} \quad \langle x | \boldsymbol{A}^\dagger | x \rangle = \lambda^\dagger \tag{2.37}
$$

then

$$
\langle x | \boldsymbol{A} | x \rangle - \langle x | \boldsymbol{A}^\dagger | x \rangle = \langle x | \boldsymbol{A} - \boldsymbol{A}^\dagger | x \rangle = \lambda - \lambda^\dagger = 0 \quad \text{if} \quad \lambda \in \Re
$$

$$
\Rightarrow \boldsymbol{A} = \boldsymbol{A}^\dagger \tag{2.38}
$$

where $\boldsymbol{A}$ is called the **Hermition Matrix**. Because the energy E is real, **the Hamiltonian is also a Hermition matrix**. From Copenhagen interpretation of quantum mechanics, the square of the coefficient c_i is the probability having the ith state and the total probability of all possible outcomes has to be one

$$
\begin{aligned}
\langle \psi | \psi \rangle &= \sum_{ij} c_i^* \langle \phi_i | c_j | \phi_j \rangle \\
&= \sum_{ij} c_i^* c_j \langle \phi_i | \phi_j \rangle = \sum_i |c_i|^2 \\
&= \text{Trace} \begin{bmatrix} |c_1|^2 & 0 & \cdots \\ 0 & |c_2|^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = 1
\end{aligned} \tag{2.39}
$$

where each element on the digonal of the matrix is the probability at such state. The momentum operator $\boldsymbol{p}$ is[5]

$$
\boldsymbol{p} = -i\hbar \frac{\partial}{\partial x} \tag{2.40}
$$

---

[5]D. J. Griffiths, Introduction to Quantum Mechanics, Chapter 1.

Thus, the expectation value of the momentum is[6]

$$\langle \boldsymbol{p} \rangle = \langle \psi | \boldsymbol{p} | \psi \rangle = \sum_{ij} c_i^* c_j \underbrace{\langle \phi_i | \boldsymbol{p} | \phi_j \rangle}_{=\delta_{ij} p_{ij}}$$

$$= \text{Trace} \left(\rho^{\text{eigenbasis}} p^{eigenbasis} \right) \tag{2.41}$$

where $\rho^{\text{eigenbasis}}$ is the eigenvalue of the density matrix[7], which is more generally defined as

$$\rho = \sum_i \rho_i |\phi_i\rangle\langle\phi_i| \tag{2.42}$$

where

$$\rho_i = \frac{n_i}{N} = |c_i|^2 = |\phi_i\rangle\langle\phi_i| \quad \text{and} \quad \sum_i \rho_i = 1 \tag{2.43}$$

n_i is the number of particles in the ith state and N is the total number of particles in a system. Check consistency of density matrix:

$$n = \int_{-\infty}^{\infty} \frac{\gamma_1 f_1 + \gamma_2 f_2}{\gamma_1 + \gamma_2} D(E) \, dE \tag{2.44}$$

Assume $\gamma_1 = \gamma_2$ and $f_1 = f_2 = f_0(E - \mu)$, we have[8]

$$n = \int_{-\infty}^{\infty} f_0(E - \mu) D(E) \, dE$$

$$= \sum_{\varepsilon_i} \int_{-\infty}^{\infty} f_0(E - \mu) \delta(E - \varepsilon_i) \, dE$$

$$= \sum_{\varepsilon_i} f_0(E - \mu) = \sum_i n_i = N\langle\psi|\psi\rangle = N \tag{2.45}$$

2.2 Bloch's Theorem

If we want to describe the electron wavefunction in a solid, we need to consider the potential V experienced by the electrons. Let's consider an one-dimensional lattice as shown in Fig. 2.1. Because there is no way to distinguish the electron wavefunction at each site, the Schrodinger equations of electrons at each site can be written as

$$H|\psi\left(\vec{r} + \vec{a}\right)\rangle = E|\psi\left(\vec{r} + \vec{a}\right)\rangle \tag{2.46}$$

$$H|\psi\left(\vec{r} + 2\vec{a}\right)\rangle = E|\psi\left(\vec{r} + 2\vec{a}\right)\rangle \tag{2.47}$$

---

[6] $\langle x|\boldsymbol{p}|x'\rangle = -i\hbar\delta'(x - x')$. See R. Shankar, Principles of Quantum Mechanics, second edition, page 64-64.

[7] It can be interpreted as the probability of having electrons in a given state.

[8] There is a confusion here. The unit of the density of states here is **per unit energy**. However, in many semiconductor textbooks, the density of states is in unit of **per unit volume (or area or length) per unit energy**. Thus, n here is the total number of electrons (carriers) instead of density.
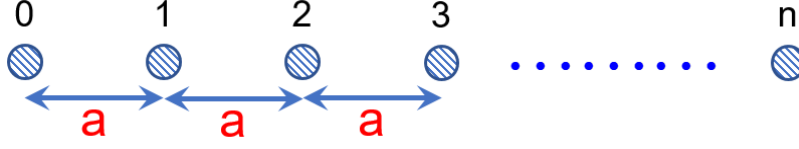
Figure 2.1: An one-dimensional lattice with lattice constant of a.

and so on. However, above equations imply infinite degeneracy in a solid which is not physical. Therefore, between each wavefunction there should be a scalar multiplier. In other words,

$$|\psi\left(\vec{r} + 2\vec{a}\right)\rangle = A|\psi\left(\vec{r} + \vec{a}\right)\rangle \tag{2.48}$$

where A is a scalar. Note that

$$\langle\psi\left(\vec{r} + 2\vec{a}\right)|\psi\left(\vec{r} + 2\vec{a}\right)\rangle = A^*A\langle\psi\left(\vec{r} + \vec{a}\right)|\psi\left(\vec{r} + \vec{a}\right)\rangle$$

$$\Rightarrow |A|^2 = 1 \Rightarrow A = e^{i\phi}$$

Thus,

$$|\psi\left(\vec{r} + n\vec{a}\right)\rangle = e^{in\phi}|\psi\left(\vec{r}\right)\rangle$$

or

$$\boxed{|\psi\left(\vec{r} + \vec{L}\right)\rangle = e^{i\vec{k}\cdot\vec{L}}|\psi\left(\vec{r}\right)\rangle} \tag{2.49}$$

where $L(=na)$ is the length of the one-dimensional lattice and $k\left(=\frac{2\pi}{a}\right)$ is the wave number. The Equation (2.49) is called the **Bloch's Theorem**. In general, the Equation (2.49) can be further expressed, for example, in three-dimensional lattice, and the relation between the wave number and the lattice size follows the Fourier transformation.

$$u\left(\vec{r}\right) = \sum_g A_g e^{i\vec{g}\cdot\vec{r}} \tag{2.50}$$

where the space spanned by $\vec{g}$ is called the **reciprocal space** [see Fig. 2.2(b)], and $\vec{g}$ in three-dimensional lattice is

$$\vec{g} = \frac{2\pi}{a}n_1\hat{x} + \frac{2\pi}{b}n_2\hat{y} + \frac{2\pi}{c}n_3\hat{z}$$
$$= n_1\vec{b_1} + n_2\vec{b_2} + n_2\vec{b_2} \tag{2.51}$$

where a, b, and c are the lattice constant of x, y, and z directions, n_i's are the integers, and $\vec{b_i}$'s are

$$\vec{b_1} \equiv \frac{2\pi}{a}\hat{x}, \quad \vec{b_2} \equiv \frac{2\pi}{b}\hat{y}, \quad \text{and} \quad \vec{b_3} \equiv \frac{2\pi}{c}\hat{z} \tag{2.52}$$

The Equation (2.52) can be generalized

$$\vec{b_1} = 2\pi \frac{\vec{a_2} \times \vec{a_3}}{\vec{a_1} \cdot \left(\vec{a_2} \times \vec{a_3}\right)}$$

$$\vec{b_2} = 2\pi \frac{\vec{a_3} \times \vec{a_1}}{\vec{a_1} \cdot \left(\vec{a_2} \times \vec{a_3}\right)}$$

$$\vec{b_3} = 2\pi \frac{\vec{a_1} \times \vec{a_2}}{\vec{a_1} \cdot \left(\vec{a_2} \times \vec{a_3}\right)} \tag{2.53}$$

where $\vec{a_i}$'s and $\vec{b_i}$'s are the basis of the real space and reciprocal space, and the denominator is the volume of the unit cell in the real space. Suppose we have a vector constructed by the basis of the real space

$$\vec{l} = ma\vec{a_1} + nb\vec{a_2} + oc\vec{a_3} \tag{2.54}$$

where m, n, and o are the integers. Then,

$$\vec{g} \cdot \vec{l} = \left(n_1\vec{b_1} + n_2\vec{b_2} + n_2\vec{b_2}\right) \cdot \left(ma\vec{a_1} + nb\vec{a_2} + oc\vec{a_3}\right)$$

$$= 2\pi \left(n_1 m + n_2 n + n_3 o\right) \equiv 2\pi N \tag{2.55}$$

This implies that the wavefunction will repeat after a certain translation. If we define $\vec{k} = \vec{k'} + \vec{g}$, the Equation (2.49) becomes

$$\psi\left(\vec{r} + \vec{L}\right) = e^{i\vec{k}\cdot\vec{L}}\psi\left(\vec{r}\right) = e^{i\left(\vec{k'}+\vec{g}\right)\cdot\vec{L}}\psi\left(\vec{r}\right)$$

$$= \underbrace{e^{i(2\pi)N}}_{=1} e^{i\vec{k'}\cdot\vec{L}}\psi\left(\vec{r}\right) \tag{2.56}$$

Since there is no way to distinguish k and k', the wavefunction will repeat after shift k by $2\pi/a$. Within the region where the wavefunction does not repeat, it is called the **Brillouin zone** as shown in Fig. 2.2(a). A common used method to obtain the Brillouin zone is to draw the **Wigner Seitz cell** which bisects each nearest neighbor as shown in Fig. 2.2(b).

2.3 Band Structure

2.3.1 One-Dimensional Monatomic Lattice

Suppose we have an one-dimensional lattice shown in Fig. 2.1 and assume that the overlap of wavefunctions only presents between nearest neighbor atoms. The elements in Hamiltonian matrix are

$$H_{ii} = \left\langle\psi_i\middle|\boldsymbol{H}\middle|\psi_i\right\rangle = t_0 \tag{2.57}$$
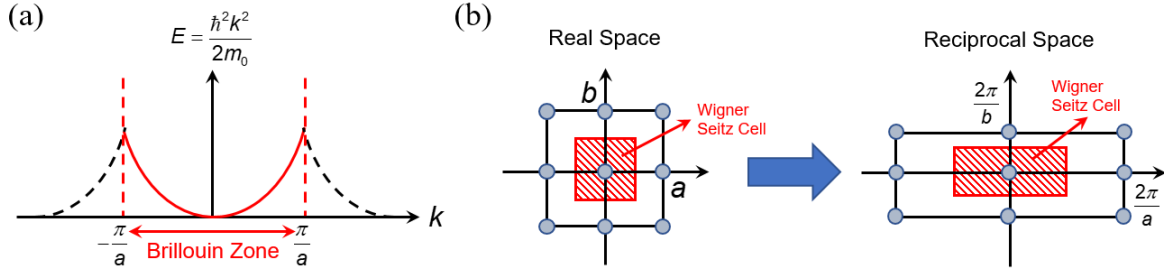
Figure 2.2: (a) Dispersion relation $(E - k)$ of an one-dimensional lattice. (b) An two-dimensional lattice in real space and reciprocal space.

$$H_{i(i\pm1)} = \langle \psi_i | \boldsymbol{H} | \psi_{i(i\pm1)} \rangle = t_1 \tag{2.58}$$

The full matrix looks like

$$\boldsymbol{H} = \begin{bmatrix} t_0 & t_1 & 0 & 0 & 0 & \dots \\ t_1^* & t_0 & t_1 & 0 & 0 & \dots \\ 0 & t_1^* & t_0 & t_1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2.59}$$

The Schrodinger equation becomes[9]

$$\boldsymbol{H}\boldsymbol{\psi} = \begin{bmatrix} t_0 & t_1 & 0 & 0 & 0 & \dots \\ t_1^* & t_0 & t_1 & 0 & 0 & \dots \\ 0 & t_1^* & t_0 & t_1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{bmatrix} = E \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{bmatrix} \tag{2.60}$$

At the nth row:

$$t_0\psi_n + t_1\psi_{n+1} + t_1^*\psi_{n-1} = E\psi_n \tag{2.61}$$

and from Block's theorem

$$\psi_n = \psi_0 e^{ikna} \tag{2.62}$$

Thus, we have

$$E = t_0 + t_1 e^{ika} + t_1^* e^{-ika} \xrightarrow{\text{if } t_1 \text{ is real}} t_0 + 2t_1 \cos ka \tag{2.63}$$

If $ka \ll 1$ and $t_1 < 0$, the Equation (2.63) can be expanded by using Taylor series.

$$\begin{aligned} E &\simeq t_0 + 2t_1 \left[1 - \frac{(ka)^2}{2} + \dots \right] \\ &= t_0 + 2t_1 - \left(t_1 a^2 \right) k^2 \\ &= t_0 - 2|t_1| + \left(|t_1| a^2 \right) k^2 \end{aligned} \tag{2.64}$$

---

[9]Each ψ represents the wavefunction of an electron at each lattice site instead of the basis function.
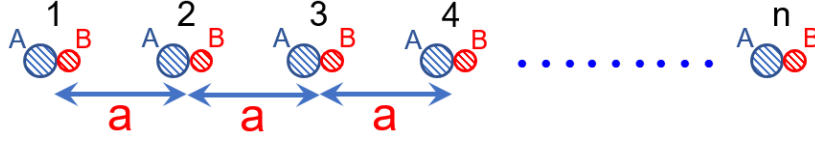
Figure 2.3: An one-dimensional diatomic (A and B) lattice with lattice constant of a.

Define the effective mass m^*

$$m^* = \frac{\hbar^2}{2a^2|t_1|} \tag{2.65}$$

The Equation (2.64) becomes

$$E = E' + \frac{\hbar^2}{2m^*}k^2 \tag{2.66}$$

where $E' \equiv t_0 - 2|t_1|$. When $|t_1|$ is large, the coupling between the nearest neighbor is strong so that the electron is "de-localized[10]." The effective mass approximation is usually valid for the transport phenomenon in the commonly used semiconductors since the electrons involving the transport are near the bottom of the band (within 1eV). However, when solving the binding strength between the nearest neighbors, this approximation may not be good.

2.3.2 One-Dimensional Diatomic Lattice

The diatomic lattice contains two atoms in a lattice site as shown in Fig. 2.3. We assume that the wavefunction overlap only happens between the nearest neighbor atoms and atoms on site, i.e., A_1 and B_1 or B_1 and A_2. Based on this assumption, the Hamiltonian can be expressed as

		1	1	2	2	3	3	
		A	B	A	B	A	B	
1	A	t_0	t_1	0	0	0	0	...
1	B	t_1^*	t_0	t_2	0	0	0	...
2	A	0	t_2^*	t_0	t_1	0	0	...
2	B	0	0	t_1^*	t_0	t_2	0	...
3	A	0	0	0	t_2^*	t_0	t_1	...
3	B	0	0	0	0	t_1^*	t_0	...
		⋮	⋮	⋮	⋮	⋮	⋮	⋱

At the nth row[11], the Schrodinger equation gives

$$[H]_{n,n}\,[\psi_0]\,e^{ikna} + [H]_{n,n+1}\,[\psi_0]\,e^{ikna}e^{ika} + [H]_{n,n-1}\,[\psi_0]\,e^{ikna}e^{-ika} = E\,[\psi_0]\,e^{ikna} \tag{2.67}$$

---

[10]The electron can move freely through the lattice sites.

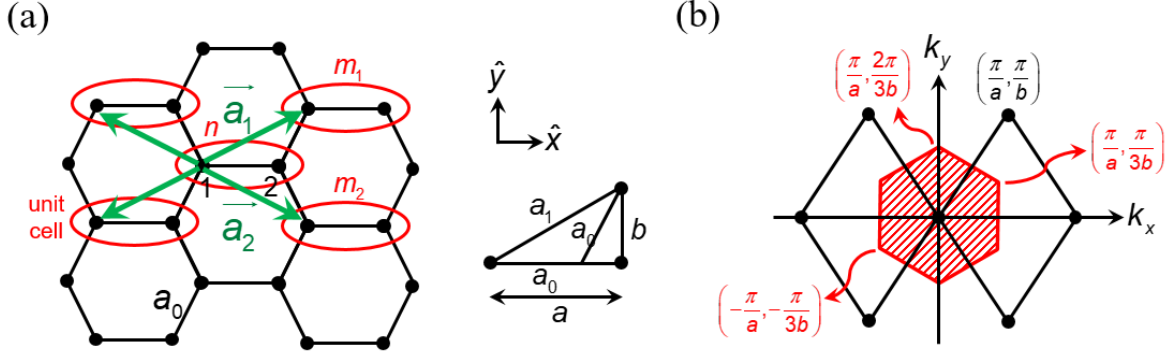[11]Each row represents each lattice site.

(a)

(b)

Figure 2.4: Graphene in (a) real space (lattice constant is a_0) and (b) reciprocal space. The green arrows in (a) are the translation vectors ($\vec{a_1}$ and $\vec{a_2}$) and two atoms circled by the red line form the unit cell. The red hexagon is the Wigner Seitz cell (Brillouin zone).

$$\Rightarrow \begin{bmatrix} t_0 & t_1 + t_2 e^{ika} \\ t_1^* + t_2^* e^{-ika} & t_0 \end{bmatrix} \begin{bmatrix} \psi_{0A} \\ \psi_{0B} \end{bmatrix} = E \begin{bmatrix} \psi_{0A} \\ \psi_{0B} \end{bmatrix}$$

$$\Rightarrow E = t_0 \pm \sqrt{|t_1|^2 + |t_2|^2 + 2\mathrm{Re}\,(t_1 t_2 e^{ika})} \tag{2.68}$$

The Equation (2.68) indicates that there are two bands in this system. If considering more interactions between more basis functions, more bands would be obtained. More generally, the Equation (2.67) can be expressed as

$$\sum_m [H]_n + [H]_{n,m} e^{i\vec{k}\cdot\left(\vec{r_m}-\vec{r}\right)} \tag{2.69}$$

where the first and second terms are the Hamiltonian matrix on the lattice site and the Hamiltonian of the overlap between the nearest neighbors respectively, and m represents the nearest neighbor. The first term also can be absorbed into the second term and the Equation (2.69) becomes

$$\boxed{\sum_m [H]_{n,m} e^{i\vec{k}\cdot\left(\vec{r_m}-\vec{r}\right)} \equiv \left[h(\vec{k})\right]} \tag{2.70}$$

which is the **Fourier transformation** between the reciprocal and real spaces.

2.3.3   Graphene

Graphene consists of carbon atoms with two-dimensional hexagonal lattice as shown in Fig. 2.4. The coordinate of graphene in real space is defined in Fig. 2.4(a). Thus, the translation vectors in real space are

$$\vec{a_1} = a\hat{x} + b\hat{y} \tag{2.71}$$

$$\vec{a_2} = a\hat{x} - b\hat{y} \tag{2.72}$$

$$\vec{a_3} = \hat{z} \tag{2.73}$$

and

$$a = a_0 + a_0 \sin\frac{\pi}{6} = \frac{3a_0}{2} \tag{2.74}$$

$$b = a_0 \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}a_0 \tag{2.75}$$

Then, the translation vectors in reciprocal space are

$$\vec{b_1} = 2\pi\frac{\vec{a_2} \times \vec{a_3}}{\vec{a_1} \cdot \vec{a_2} \times \vec{a_3}} = \frac{\pi}{a}\hat{x} + \frac{\pi}{b}\hat{y} \tag{2.76}$$

$$\vec{b_2} = \frac{\pi}{a}\hat{x} - \frac{\pi}{b}\hat{y} \tag{2.77}$$

Due to weaker coupling of p_z orbitals between nearest neighbors than that of $p_{x,y}$, the energy split is smaller (the energy states are closer to the Fermi level) so that the wavefunction overlap of p_z orbitals is considered. We also assume that only the wavefunction overlap of the nearest neighbors does matter. Thus, the Hamiltonian matrix of the nth unit cell [see Fig. 2.4(a)] can be expressed using the Equation (2.70)

$$[H]_{n,n} + [H]_{n,m_1} e^{i\vec{k}\cdot\vec{a_1}} + [H]_{n,m_2} e^{i\vec{k}\cdot\vec{a_2}} + [H]_{n,m_1} e^{-i\vec{k}\cdot\vec{a_1}} + [H]_{n,m_2} e^{-i\vec{k}\cdot\vec{a_2}} \tag{2.78}$$

where

$$[H]_{n,n} = \begin{bmatrix} t_0 & t \\ t & t_0 \end{bmatrix}, \quad [H]_{n,m_1} = \begin{bmatrix} 0 & 0 \\ t & 0 \end{bmatrix}, \quad [H]_{n,m_2} = \begin{bmatrix} 0 & 0 \\ t & 0 \end{bmatrix} \tag{2.79}$$

and

$$t_0 = \langle\psi_1|\boldsymbol{H}|\psi_1\rangle, \quad t = \langle\psi_1|\boldsymbol{H}|\psi_2\rangle, \quad t = t^* \tag{2.80}$$

Here, $\psi_{1,2}$ are the wavefunctions at the first (left side) and the second (right side) lattice sites in an unit cell. Therefore, the Equation (2.78) is
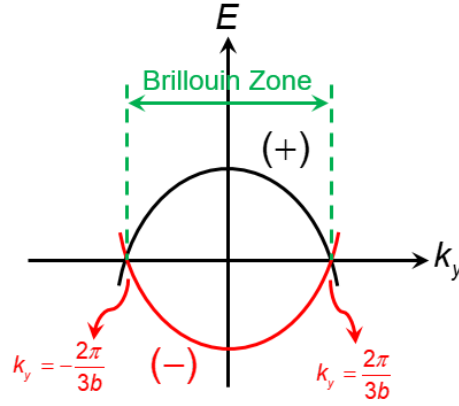
$$\left[h(\vec{k})\right] = \begin{bmatrix} t_0 & t + te^{-i\vec{k}\cdot\vec{a_1}} + te^{-i\vec{k}\cdot\vec{a_2}} \\ t + te^{i\vec{k}\cdot\vec{a_1}} + te^{i\vec{k}\cdot\vec{a_2}} & t_0 \end{bmatrix} \equiv \begin{bmatrix} t_0 & p^*(\vec{k}) \\ p(\vec{k}) & t_0 \end{bmatrix} \tag{2.81}$$

where

$$p(\vec{k}) = t + te^{i\vec{k}\cdot\vec{a_1}} + te^{i\vec{k}\cdot\vec{a_2}} = t\left[1 + 2e^{ik_x a}\cos k_y b\right] \tag{2.82}$$

The eigenvalue (energy) of the matrix (2.81) is

$$E(\vec{k}) = t_0 \pm \left|p(\vec{k})\right|$$

$$= t_0 \pm t\sqrt{1 + 4\cos^2 k_y b + 4\cos k_x a \cos k_y b} \tag{2.83}$$

Figure 2.5: E-k diagram of graphene at $k_x = 0$.

The above equation describes the dispersion relationship (E-k diagram) of graphene. Let's now consider $k_x = 0$:

$$E = t_0 \pm t \left(1 + 2\cos k_y b\right) \tag{2.84}$$

If $t_0 = 0$ is chosen, the E-k_y plot is shown in Fig 2.5. Note that $k_y = \pm\frac{2\pi}{3b}$ are the points located at the Brillouin zone boundary as shown in Fig. 2.4(b). These points are called "Dirac points." If $k_x \neq 0$, the E-k would look like a cone and the k_x and k_y satisfying $E = 0$ are $\left(\pm\frac{\pi}{a}, \pm\frac{\pi}{3b}\right)$ where are at the zone boundary. Because the two bands are crossing at the Dirac points, the graphene behaves like a metal and the number of electrons near the Fermi level ($E_F = 0$) is $2 \times 2 \times 2 = 8$[12]. Now, let's consider E near the Dirac point. The Taylor expansion of the Equation (2.83) assuming $t_0 = 0$ is

$$
\begin{aligned}
E(\vec{k}) &= \pm\left|p(\vec{k})\right| \\
&\approx \pm\left|\frac{\partial p}{\partial k_x}\Big|_{0,\pm\frac{2\pi}{3b}}(k_x - 0) + \frac{\partial p}{\partial k_y}\Big|_{0,\pm\frac{2\pi}{3b}}\left(k_y \pm \frac{2\pi}{3b}\right)\right| \\
&= \pm\left| -\frac{3}{2}a_0 t\left[ik_x + \left(k_y \mp \frac{2\pi}{3b}\right)\right]\right| \\
&= \pm\frac{3}{2}a_0 t\sqrt{k_x^2 + \left(k_y \mp \frac{2\pi}{3b}\right)^2} \\
&= \pm at\left|k\right| \quad \text{(linear dispersion relationship)}
\end{aligned}
\tag{2.85}
$$

which is different from the square dependence of k for the free electron. The definition of the effective mass $m^* = \frac{\hbar^2}{\frac{\partial^2 E}{\partial k^2}}$ at the Dirac points becomes physically meaningless where defining $p = \hbar k$ would be better. Since $p = 0$ at the Dirac points, there will be no current flow at

---

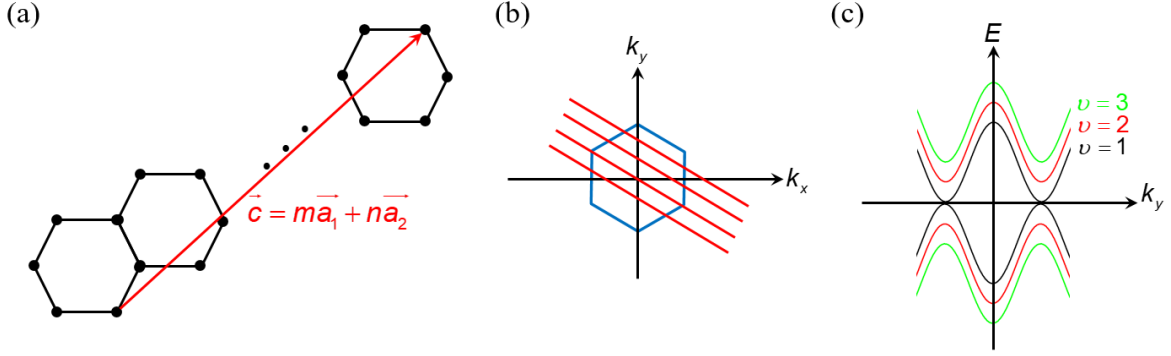[12]2's are coming from two bands, up- and down-spin, and equivalent corner in the Brillouin zone.

Figure 2.6: (a) Graphene with rolling vector $\vec{c}$. (b) Brillouin zone of graphene and possible solutions (red straight lines) of nanotube. (c) $E - k$ diagram of graphene nanotube, assuming $k_x = 0$.

that points. Around the Dirac point, the electron group velocity is

$$v = \frac{\partial \omega}{\partial k} = \frac{1}{\hbar}\frac{\partial E}{\partial k} = \frac{3}{2}\frac{a_0 t}{\hbar} = \text{constant} \tag{2.86}$$

where v is always written as v_F (Fermi velocity). Thus, the E-k relation of graphene also can be written in terms of v_F

$$E = \pm \hbar v_F |k| \tag{2.87}$$

where the electron behaves like a relativistic particle and cannot be accelerated.

2.3.4 Graphene Nanotube

Since the graphene behaves like a metal, engineering graphene to be semiconductor would be very important. One way to achieve that is to roll it into a tube as shown in Fig. 2.6. After rolling the graphene, we impose additional restriction to the lattice

$$\vec{k} \cdot \vec{c} = 2\pi \nu, \quad \nu \in Z \tag{2.88}$$

because the wavefunction will be a periodic function. The Equation (2.88) can be further written as

$$k_y = \frac{2\pi}{(m-n)\,b}\nu - \frac{(m+n)\,a}{(m-n)\,b}k_x \tag{2.89}$$

which represents straight lines shown in Fig. 2.6(b) depending on ν if m and n are known. Because the corners of the Brillouin zone [see Fig. 2.6(b)] correspond to the Dirac points, the graphene nanotube exhibits the metallic nature if the straight line crosses the corners. Therefore, whether the graphene nanotube behaves like a metal or a semiconductor would depend on how it is rolled (various combinations of m and n).
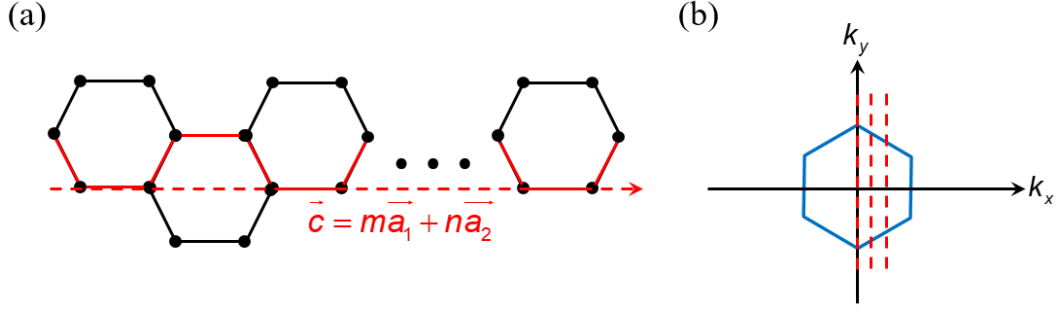
(a)   (b)

Figure 2.7: (a) Illustration of rolling an armchair nanotube. (b) Brillouin zone and possible solutions (straight lines) of armchair nanotube.

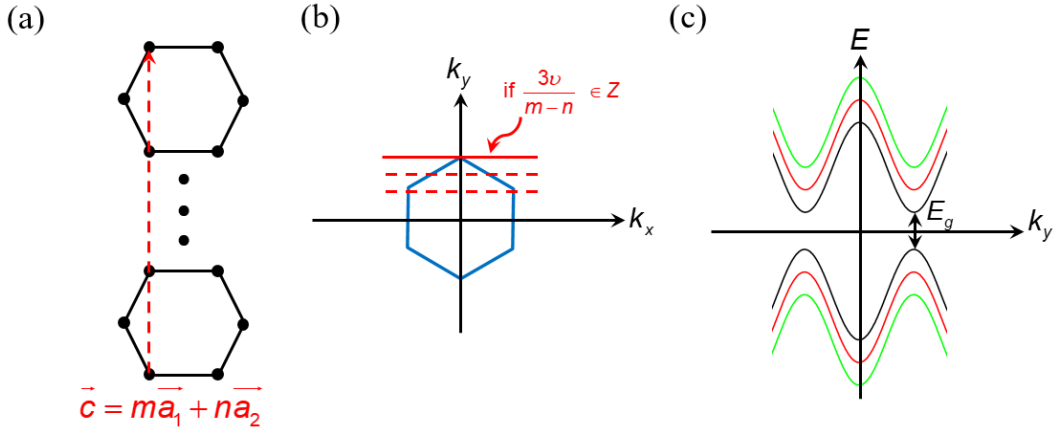(a)   (b)   (c)

Figure 2.8: (a) Illustration of rolling a zigzag nanotube. (b) Brillouin zone and possible solutions (straight lines) of zigzag nanotube. (c) $E - k_y$ diagram when $(m - n)$ is not an integer multiple of 3.

Armchair Graphene Nanotube

If the graphene is rolled only in x-direction, it's called armchair nanotube as shown in Fig. 2.7. Because now the wavefunction is restricted in x-direction, the k_y is zero and thus the Equation (2.89) becomes

$$k_x = \frac{2\pi}{(m + n)\, a}\nu \tag{2.90}$$

Since $k_x = 0$ is a valid solution, the armchair nanotube is always metallic.

Zigzag Graphene Nanotube

If the graphene is rolled only in y-direction, it's called zigzag nanotube as shown in Fig. 2.8. Because now the wavefunction is restricted in y-direction, the k_x is zero and thus the Equation (2.89) becomes

$$k_y = \frac{2\pi}{3b} \frac{3\nu}{m - n} \tag{2.91}$$

If $\frac{3\nu}{m-n}$ is an integer, the zigzag nanotube will become metallic since k_y crosses the corner of Brillouin zone [see Fig. 2.8(b)]. If $\frac{3\nu}{m-n}$ is not an integer, the zigzag nanotube behaves like a semiconductor and there is a bandgap E_g as shown in Fig. 2.8(c). The bandgap can be evaluated using the Equation (2.85)

$$
\begin{aligned}
E_g &= 2 \times \frac{3}{2} a_0 t \sqrt{k_x^2 + \left(k_y \mp \frac{2\pi}{3b}\right)^2}\Big|_{k_x=0,k_y=\frac{2\pi}{3b}\frac{3\nu}{m-n}} \\
&= 3 a_0 t \frac{3}{m-n}\left(\nu \mp \frac{m-n}{3}\right)\frac{2\pi}{3b} \\
&\xrightarrow{minimum} 3 a_0 t \frac{3}{m-n}\frac{1}{3}\frac{2\pi}{3b} = \frac{2 a_0 t \pi}{(m-n)\,b}
\end{aligned}
\tag{2.92}
$$

If the diameter of the zigzag nanotube is d, the bandgap is

$$
E_g = \frac{2 a_0 t}{d} = \frac{0.8}{d\ (\mathrm{nm})}\ \mathrm{eV}
\tag{2.93}
$$

where $\pi d = (m-n)b$. The Equation (2.93) is validated by the experimental results. However, if $d < 1$ nm, it is not accurate (but the trend is still correct) due to the presence of strain.