Quick review of Lecture 5

In the last lecture, we approximate the potential field as rectangular wells in the crystal. From periodicity, the Bloch theorem gives additional equations

$$\Psi[x + (a+b)n] = \Psi(x)e^{ikn(a+b)},$$

which is used to determine the coefficients in $\Psi = Ae^{-iKx} + Be^{iKx}$. We have also determined the value of the wavevector k in the Bloch theorem, using the **Born-von Karman** periodic boundary condition $\Psi[x + N(a+b)] = \Psi(x)$. This yields allowed k values as

$$k = \frac{2\pi n}{N(a+b)} = \frac{2\pi n}{L} \quad (n=0, \pm 1, \pm 2, ...),$$

where L is the length of the crystal.

For each k_n, there are two quantum states denoted by $\Psi_{k_n,s}$ (spin up, spin down). When n goes from 1 to N, k_n varies from 0 to $2\pi/a$. Therefore, in the following E-k figure we totally have 2N quantum states $\Psi_{k_n,s}$.

Note: (1) For big crystals, N is very big (on the magnitude of 10^{23}) and $\Delta k = \frac{2\pi}{L}$ is also small. The following curve can be regarded as quasi-continuous. (2) The **Born-von Karman** periodic boundary condition is no longer valid when N becomes very small in nanomaterials.
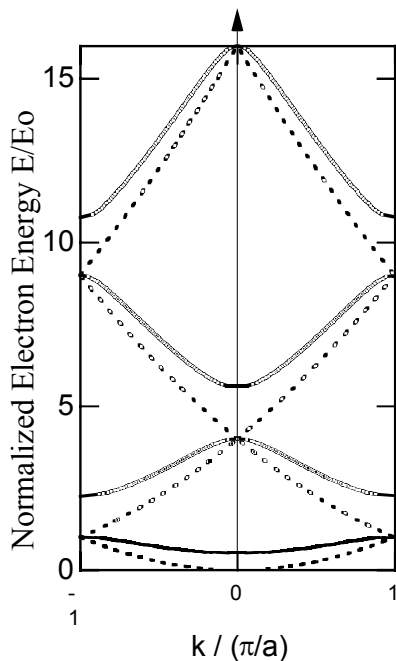
3.1.3 Consequences of solid energy levels we just obtained:
(1) Electrons wave function extends through the whole crystal, they belong to all atoms collectively.

Recall the splitting of waves discussed in last lecture. According to Pauli's exclusion principle, the wavefunctions of adjacent wells cannot overlap. A continuous wave extending through the whole structure will be formed in this situation. The wavefunction no longer corresponds to an individual atom.
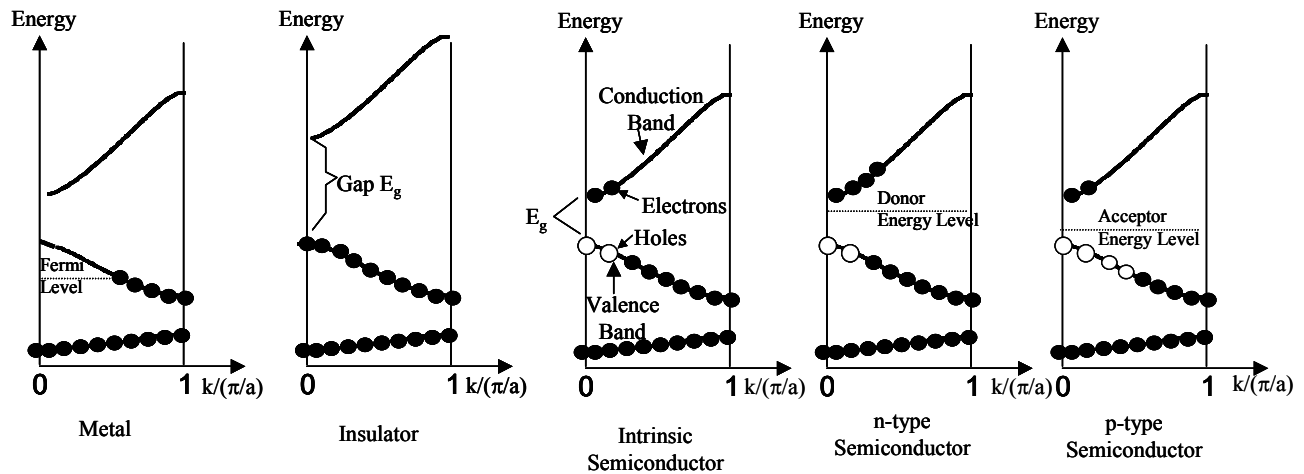

NOT allowed by Pauli exclusion principle

Continuous wavefunction across the structure
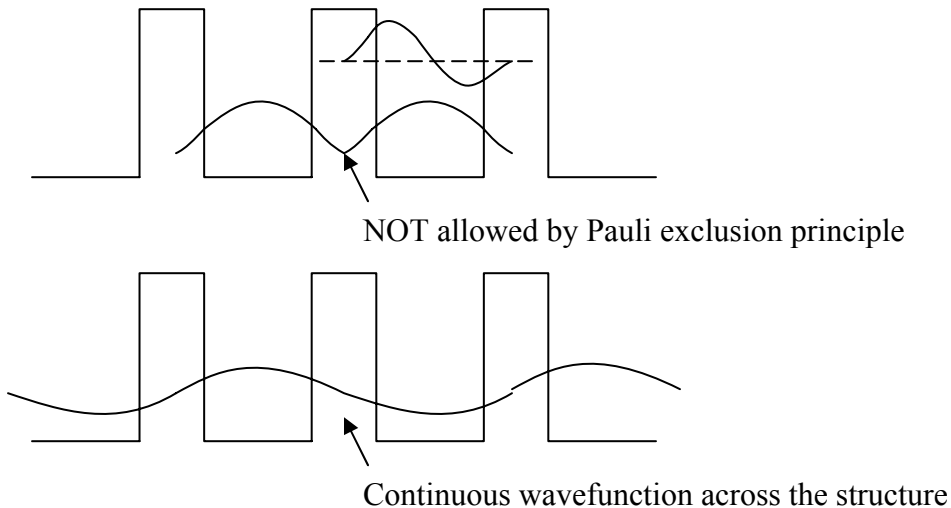
Similar argument also exists for atomic energy levels. When two atoms become closer, the overlap of electron wave functions will cause band split.

-13.6 eV/n^2

-13.6 eV
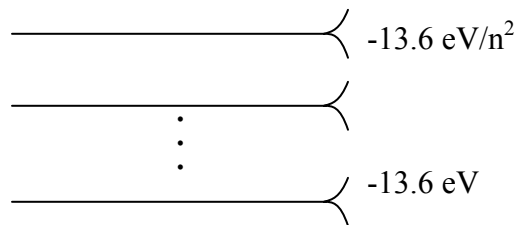
(2) Filling of electrons
As mentioned before, in E-k figure every band (k changes for $2\pi/a$) can accommodate 2N quantum states. At zero temperature, the filling rule for the electrons is that they always fill the lowest energy level first, as required by thermodynamics. If one atom only has one electron, the band is half filled since there are only N valence electrons in this case, as shown in the next figure. The topmost energy level that is filled with electrons at zero Kelvin is called the **Fermi level**. The electron energy and momentum can be changed (almost) continuously within the same band because the separation between successive energy levels is small. Thus, these electrons can flow freely, making the materials good electrical conductors, which is the case for metals.

Metal — Energy vs k/(π/a), with Fermi Level indicated

Insulator — Energy vs k/(π/a), with Gap E_g

Intrinsic Semiconductor — Energy vs k/(π/a), with Conduction Band, Electrons, E_g, Holes, Valence Band

n-type Semiconductor — Energy vs k/(π/a), with Donor Energy Level

**p-type Semiconductor** — Energy vs k/(π/a), with Acceptor Energy Level

If the valence electrons exactly fill one or more bands, leaving others empty, the crystal will be an insulator at zero temperature and can be an insulator or a semiconductor at other temperatures depending on the value of the energy gap. If a filled band is separated by a large energy gap (>3 eV) from the next higher band, one cannot change the energy and the momentum of an electron in the filled band easily, that is, these electrons cannot move freely and the materials are insulators. A semiconductor is essentially similar to an insulator. The difference between them is that the gap between the filled and the empty bands for a semiconductor are not so large (<3 eV), so that some electrons have enough thermal energy (at room temperature k_BT = 0.026 eV) to jump across the gap to the empty band above (called conduction band), and these electrons can conduct electricity (these are called intrinsic semiconductors). The unoccupied states left behind also leave room for the electrons in the original band (called valence band) to move. It turns out that the description of the motion of these electrons is equivalent to thinking that the vacant states move as positive electrons, or holes. The energy of these holes is a minimum at the peak in the valence band and increases as the electron energy becomes more negative.

Note: (1) Here we only talk about 1D case, for 3D structure we have different band shapes in different direction and the wavefunction changes to $\Psi(k_x, k_y, k_z)$. Later we will talk about the electron filling in a 3D structure. (2) The shape of bands can be affected by heavily doping. (3) The lowest band starts from a nonzero energy, which is the consequence of the uncertainty principle.

Impurities are added to most semiconductors and these impurities have energy levels somewhere within the band gap, some are close to the bottom edge of the conduction band or top edge of the valence band (or band edge). The electrons in the impurity levels can be thermally excited to the conduction band if their level is close to the bottom of the conduction band, creating more electrons in the semiconductor than holes. Such semiconductors are called n-type and the impurities are called donors. A typical example is P. Similarly, if the impurity energy level is close the valence band edge, electrons in

the valence band can be excited to the impurity levels, leaving more empty states or holes behind. Such semiconductors are called p-type and the impurities are called acceptors. For instance, a B atom has three outer core electrons; it will catch one more electron from Si atoms.
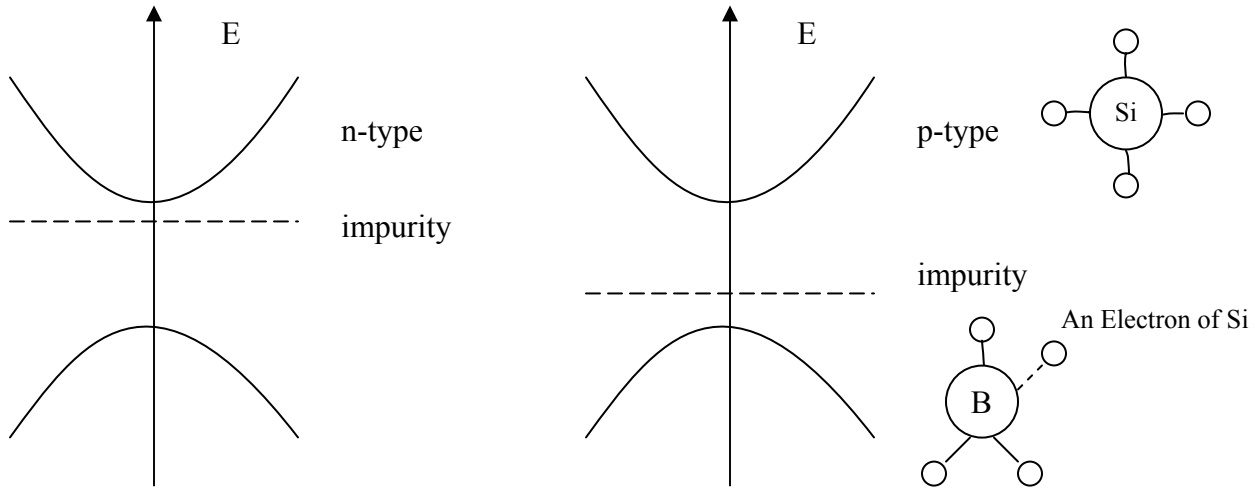
In semiconductors, the moving charge carriers normally are near the minima or maxima of a band, where $\dfrac{\partial E}{\partial k} = 0$. Taylor's expansion gives

$$E(k) = E(k_m) + \frac{\partial E}{\partial k}\Big|_{k_m}(k - k_m) + \frac{1}{2}\frac{\partial^2 E}{\partial k^2}\Big|_{k_m}(k - k_m)^2$$

$$= E(k_m) + \frac{1}{2}\frac{\partial^2 E}{\partial k^2}\Big|_{k_m}(k - k_m)^2$$

$$= E(k_m) + \frac{1}{2}\frac{\hbar^2}{m^*}(k - k_m)^2,$$

where the effective mass is defined as $m^* = \dfrac{\hbar^2}{(\partial^2 E / \partial k^2)|_{k_m}}$. In differential geometry, the term $1/\partial^2 E / \partial k^2$ is just the curvature. Thus, effective mass is proportional to local curvature at band maxima or minima. For electrons close to the minima of the conduction band, we have

$$E - E_c = \frac{\hbar^2 k^2}{2m^*},$$

from which we can clearly see the meaning of effective mass by comparing with the energy of free electrons

$$E = \hbar^2 k^2 / 2m.$$

In the above equation, $\hbar k$ represents the momentum of the free electron. The momentum of an electron in the crystal, however, should be calculated from the wavefunction using the momentum operator $-i\hbar\nabla\Psi$. Such a calculation would show that $\hbar k$ is not the momentum of the electron. Nevertheless, in many ways $\hbar k$ for a periodic potential behaves as the momentum of a free electron and thus it is called the **crystal momentum**.

3.2 Different directions of a crystal, different E(k) relation
In the real crystal, the band shapes normally differ a lot in different directions. In the following figure, the periods are a, $\sqrt{2}a$, and $\sqrt{3}a$ respectively in three directions.
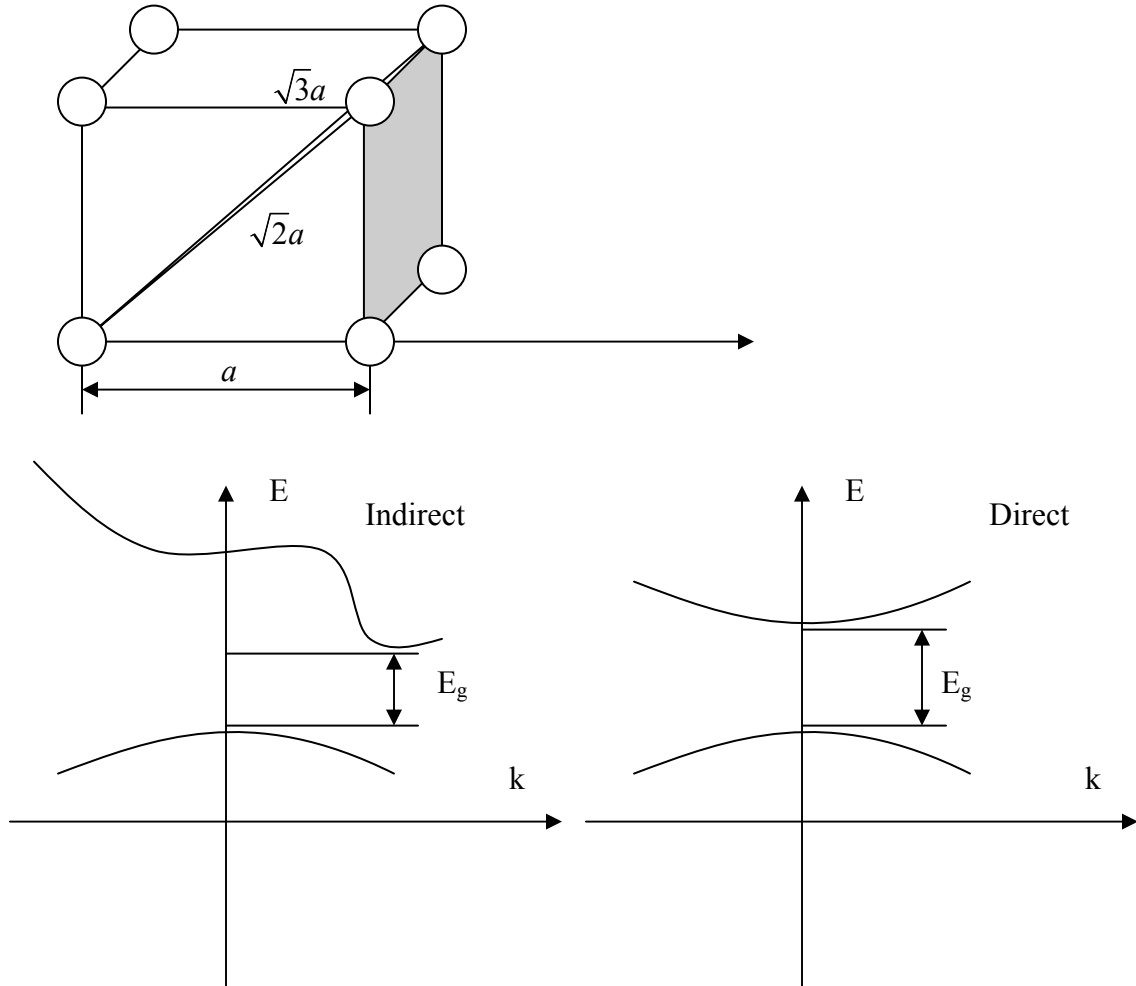
In above figures, we demonstrate the idea of direct and indirect band gaps. In a direct band structure, both the minimum in the conduction band and the maximum in the valence band occur at the same location of k (=0 for the example given). A good example for direct band gap is GaAs with E_g =1.42 eV, while Si is indirect gap semiconductor. Direct gap semiconductors are used in lasers, while Si is for microelectronics.
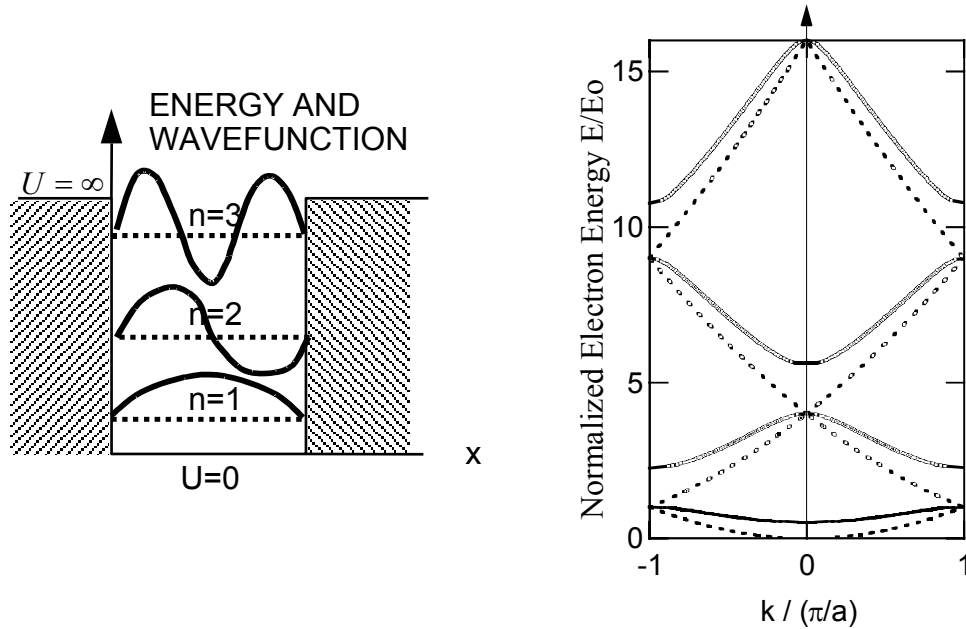
3.3 Lattice vibration and phonons
We have considered a harmonic oscillator for the H_2 molecule. The energy is expressed as
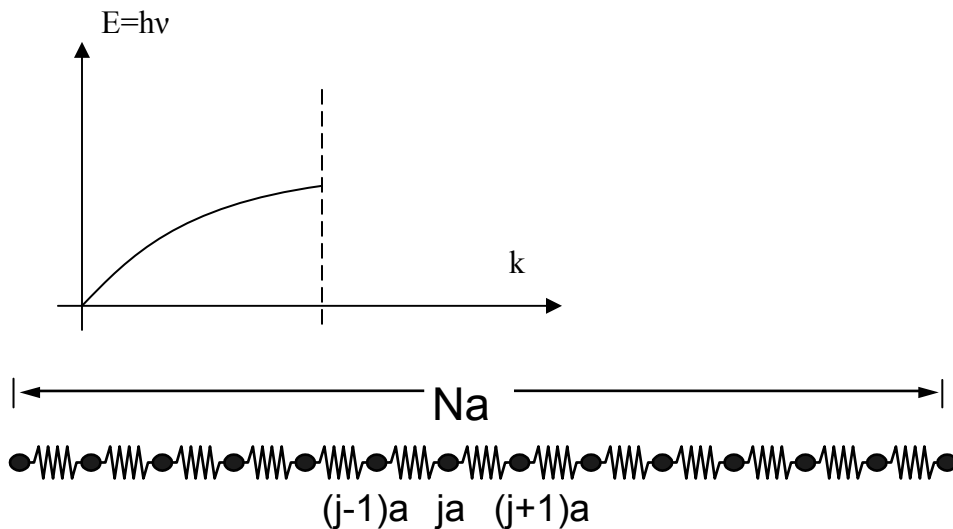
$$E_n = h\nu(n + \frac{1}{2}); \nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \quad (n=0,1,2\ldots)$$

Now let us think about N atoms in a one-dimensional chain. First recall the following cases. In the first figure, the energy is quantized, while band gap appears in the second figure.

Similarly, we can anticipate the single energy level of a diatomic molecule will split into a band in a lattice chain of atoms.

We make the following assumptions for the analysis. First, we consider only the interaction force between the nearest neighbors. Second, the interaction force between atoms is a harmonic force (which obeys Hook's law). This can be justified as we have done for harmonic oscillators. Now consider a typical atom j. The displacement of atom j from its equilibrium position x_j^o is

$$u_j = x_j - x_j^o$$

The force acting on atom j comes from two parts. One is due to the relative displacement between atom (j-1) and atom j, and the other is due to the relative displacement between atom j and (j+1). The net force is then

$$F_j = K(u_{j+1} - u_j) - K(u_j - u_{j-1}).$$

Applying Newton's second law to atom j, we obtain

$$m\frac{d^2 u_j}{dt^2} = K(u_{j+1} - u_j) - K(u_j - u_{j-1}).$$

The above equation is a special form of the differential wave equation

$$m\frac{\partial^2 u}{\partial t^2} = Ka^2 \frac{\partial^2 u}{\partial x^2}.$$

which has a solution of the form $u \propto e^{-i(\omega t - kx)}$. Since a is not very small, we cannot use this solution directly. But this leads us to guess a wave type of solution as

$$u_j = A \exp[-i(\omega t - kja)].$$

Substituting the guessed solution into the discrete equation, we get

$$-m\omega^2 = K\left[e^{ika} + e^{-ika} - 2\right]$$

or

$$\omega = 2\sqrt{\frac{K}{m}}\left|\sin\frac{ka}{2}\right|.$$