

## 2.1.5 Tight-binding approximation

In the nearly free electron approximation (NFEA), the lattice potential causes interference of electron waves in free space and opens "slits", namely energy gaps in the original continuous energy spectrum. On the other hand in the view of tight-binding approximation (TBA), the energy bands with finite widths are formed from originally discrete energy levels in spatially localized potential at each site by electron tunneling between the different sites.



Fig. 2.4 Schematic diagram of TBA. The periodic potential (blue lines) is the sum of atomic potentials represented by broken lines. In TBA, the linear combination of localized orbitals drawn by red lines should be in the form of Bloch function.

The concept is illustrated in Fig. 2.4. The Hamiltonian for single-site is written as

$$\mathscr{H}_a = \hat{T} + u, \tag{2.12}$$

where  $\hat{T}$ , u are the kinetic energy and the site-localized potential respectively. Let  $\mathbf{R}_i$  be the potision of site i, then the real-space representation of  $\mathscr{H}_a$  can be written as  $\mathscr{H}_a(\mathbf{R}_i) = \hat{T} + u(\mathbf{r} - \mathbf{R}_i)$ . If we write the normalized orthogonal eigenstates of  $\mathscr{H}_a(0)$  as  $\{\phi_n\}$ , then

$$\mathscr{H}_{a}(\boldsymbol{R}_{i})\phi_{n}(\boldsymbol{r}-\boldsymbol{R}_{i}) = \epsilon_{n}\phi_{n}(\boldsymbol{r}-\boldsymbol{R}_{i}), \qquad (2.13)$$

where n is the level index of the discrete states. We write the periodic potential obtained by overlapping of u as V(r), the total Hamiltonian for such system illustrated in Fig. 2.4 is

$$\mathscr{H} = \hat{T} + V(\boldsymbol{r}). \tag{2.14}$$

Let us consider the solution of the eigenequation of the above Hamiltonian. For simplicity we ignore the direct overlapp integral of wavefunctions  $\phi_n$  between the neighboring sites $(\langle \phi_n(\mathbf{r} - \mathbf{R}_i) | \phi_n(\mathbf{r} - \mathbf{R}_j) \rangle = \delta_{ij})$ . We also ignore the matrix elements of  $\mathscr{H}$  between the states of different n. Then the total wavefunction can be written in the linear combination of  $\phi_n$  Because all of the lattice sites are equivalent, each coefficient of the linear combination should be in the form of  $c/\sqrt{N}$  where |c| = 1, c is constant for  $\mathbf{r}$ . Furthermore, the linear combination should be written in the form of the Bloch function. From the above requirements the eigenstate should be written in the following form.

$$\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N}} \sum_{i} e^{i\boldsymbol{k}\cdot\boldsymbol{R}_{i}} \phi_{n}(\boldsymbol{r}-\boldsymbol{R}_{i}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{\sqrt{N}} \left[ \sum_{i} e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{R}_{i})} \phi_{n}(\boldsymbol{r}-\boldsymbol{R}_{i}) \right].$$
(2.15)

Since the last part of  $[\cdots]$  has the periodicity of lattice, the form of (2.15) is in the Bloch form.



Fig. 2.5 Band dispertion relation in onedimensional tight-binding approximation. In the model the hopping integral is finite only for nearest neighbor and t, in which situation cosine-shaped band structure is obtained. The insets illustrate the standing wave amplitude, phase at the band bottom and the top.

The expectation value of  $\mathscr{H}$  (energy expectation value) with  $\psi_{nk}$  is

$$\langle \psi_{n\boldsymbol{k}} | \mathscr{H} | \psi_{n\boldsymbol{k}} \rangle = N^{-1} \sum_{i,j} e^{i\boldsymbol{k} \cdot (\boldsymbol{R}_{j} - \boldsymbol{R}_{i})} \langle \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{i}) | [\hat{T}_{r} + V(\boldsymbol{r})] | \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{j}) \rangle$$

$$= N^{-1} \sum_{i,j} e^{i\boldsymbol{k} \cdot (\boldsymbol{R}_{j} - \boldsymbol{R}_{i})} \langle \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{i}) | [\hat{T}_{r} + u(\boldsymbol{r} - \boldsymbol{R}_{i}) + V(\boldsymbol{r}) - u(\boldsymbol{r} - \boldsymbol{R}_{i})] | \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{j}) \rangle$$

$$= \epsilon_{n} + N^{-1} \sum_{i,j} e^{i\boldsymbol{k} \cdot \boldsymbol{R}_{j}} \langle \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{i}) | [V(\boldsymbol{r}) - u(\boldsymbol{r} - \boldsymbol{R}_{i}))] | \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{j}) \rangle$$

$$= \epsilon_{n} + \sum_{j} e^{i\boldsymbol{k} \cdot \boldsymbol{R}_{j}} \langle \phi_{n}(\boldsymbol{r}) | [V(\boldsymbol{r}) - u(\boldsymbol{r}))] | \phi_{n}(\boldsymbol{r} - \boldsymbol{R}_{j}) \rangle.$$

$$(2.16)$$

From the second line to the third line, we have used the fact that  $\hat{T}_r + u(r - R_i)$  is the local Hamiltonian of the site *i* and ignored the overlapp integral. Then to the fourth line, because of the infinite integration over *r* is taken, we have shhifted the origin to each  $R_i$  for the integration, which results in simply *N*-times of the result and the normalization is considered. So if we write the difference between the lattice potential V(r) and the localized potential u(r) as  $v(r) \equiv V(r) - u(r)$ ,

$$E_n(\mathbf{k}) = \epsilon_n + \langle \phi_n(\mathbf{r}) | v(\mathbf{r}) | \phi_n(\mathbf{r}) \rangle - \sum_{\mathbf{R}_j \neq 0} e^{i\mathbf{k} \cdot \mathbf{R}_j} t_n(\mathbf{R}_j), \qquad (2.17)$$

where the **hopping integral**  $t_n(\mathbf{R}_j)$  is defined as

$$t_n(\mathbf{R}_j) \equiv -\langle \phi_n(\mathbf{r}) | v(\mathbf{r}) | \phi_n(\mathbf{r} - \mathbf{R}_j) \rangle.$$
(2.18)

Let us restrict ourselves to one-dimensional systems. For simplicity, we assume the hopping integral has a finite value t only for the nearest neighbor sites. Then the sum in (2.17) is only for  $R_j = \pm a$ , where a is the lattice constant. The result gives

$$E_n(k) = \epsilon_n - \alpha_n - t(e^{ika} + e^{-ika}) = \epsilon_n - \alpha_n - 2t\cos ka.$$
(2.19)

The cosine band in (2.19) is common for one-dimensional systems.  $\alpha_n \equiv -\langle \phi_n(r) | v(r) | \phi_n(r) \rangle$  is the on-site energy shift due to the change from the localized potential to the crystal potential and called **crystal field** contribution.

Figure 2.5 shows the cosine band thus obtained. In the case of NFEA, unperturbed is the band bottom k = 0 and the perturbation is stronger with coming close to the edge of Brillouin zone. In the case of TBA, the starting point is the band center. When t = 0 the band is flat, and with increasing t it broadens to  $\pm 2t$ .

At the band bottom, as we know from (2.19), the eigenstate is in the mode of standing wave. As illustrated in one of the insets, the localized wavefunctions on the sites are synchronously summed up wih the same phase at the bottom. On the other hand, they are summed up with opposite phase for neighboring site at the band top to be standing wave with higher energy. These features are the same as those in a double well potenntial with a tunneling matrix element t between them. The bottom corresponds to the **bonding orbital** while the top to the **anti-bonding orbital**.

# 2.2 Band structure measurement, calculation

The importance of the energy band structure, that is, the energy dispersion relation in the first Brillouin zone, to materials science and engineering cannot be overstated. The views from the two extremes in the previous section are useful for understanding the band concept and for obtaining a general physical perspective on the phenomena that occur there. For quantitative analysis of experimental data or for selection of materials to design devices with desired characteristics, precise band parameters are required.

In the band structure, to analyze the optical response and hot electron characteristics, the dispersion in a wide range of crystal wavenumbers is important. Even in the optical response, it is necessary to know the dispersion of the band edge precisely for the light emission from the band edge, problems such as excitons, and electrical conduction to a low electric field. Let's call the former "global band structure" and the latter "band edge structure". Here are some typical experimental and theoretical methods to obtain for each. Appendix 2A introduces one of the ab-initio calculation methods that calculate without directly relying on the measured values of the experiment.

## 2.2.1 Measurement of global band structure: angle-resolved photoemission spectroscopy

In recent years, the resolution of photoemission spectroscopy has improved remarkably, and angle-resolved photoemission spectroscopy (ARPES) is now a means for directly obtaining a global band structure.

The right figure shows the concept of photoemission spectroscopy. Let  $E^{\nu}$  be the energy of photo-electron measured from the vacuume level,  $E_{\rm B}$  the binding energy of electrons measured from the Fermi level  $E_{\rm F}$ ,  $h\nu$  the energy of incoming light,  $\phi$  the work function of the specimen. Then in the light absorption process, the energy conservation law

$$E^{\nu} = h\nu - \phi - E_{\rm B} \tag{2.20}$$

holds. In the experiments,  $E_{kin} = \phi + E^{\nu}$  is measured and the relation is simply written as

$$E_{\rm B} = h\nu - E_{\rm kin}, \qquad (2.21)$$



which gives the binding energy. In the above ordinary photoemission spectroscopy, because the photo-electrons are collected independent of the angles, the total density of states is measured.

In the emission of photo-electron, the sum of the parallel component of crystal wavenumber  $k_{\parallel}$  and the recprocal lattice vector G is conserved. Let  $k_i$  be the wavenumber of the initial state,  $k_f$  the wavenumber of electrons emitted into the vacuum, then

$$(\boldsymbol{k}_{\mathrm{i}} + \boldsymbol{G})_{\parallel} = \boldsymbol{k}_{\mathrm{f}\parallel}. \tag{2.22}$$

The energy conservation reads

$$E_{\rm i}(\mathbf{k}_{\rm i}) + h\nu = E_{\rm f}(\mathbf{k}_{\rm i}) = \frac{\hbar^2 k_{\rm f}^2}{2m_0} + \phi.$$
 (2.23)

If we know the workfunction  $\phi$ , the dispersion relation of the final states in the crystal  $E_{\rm f}(\mathbf{k})$ , then we can obtain the dispersion relation  $E_{\rm i}(\mathbf{k})$ . Since the energy of the final state is high, the effect of lattice potential on it is small, hence



**Fig.** 2.6 (a) Illustration o the procesure to determine the dispersion relation in ARPES measurement. (b) Band structure obtained for GaAs with the vertical emission method.

 $E_{\rm f}(\mathbf{k})$  is often replaced with that of free electron. Care should be taken that the wavenumber is still a crystal wavenumber and the energy conservation law (2.23) becomes

$$E_{\rm i}(\boldsymbol{k}_i) + h\nu = \frac{\hbar^2 (\boldsymbol{k}_{\rm i} + \boldsymbol{G})^2}{2m_0} + V_0 = \frac{\hbar^2 k_{\rm f}^2}{2m_0} + \phi.$$
(2.24)

Here, still the potential  $V_0$  for the zero-kinetic energy "free electron in the crystal" is not known. But  $V_0$  is also determined for various interpretations of experiments to be consistent.

As a simplest method among ARPES, a possible way to obtain the dispersion relation is to measure photo-electrons emitted vertically to the surface. In this case, as  $k_{\parallel} = 0$ , the energy conservation gives

$$E_{\rm i}(\boldsymbol{k}_{\rm i}) + h\nu = \frac{\hbar^2 |(\boldsymbol{k}_{\rm i} + \boldsymbol{G})_{\perp}|^2}{2m_0} + V_0 = \frac{\hbar^2 k_{\rm f}^2}{2m_0} + \phi, \qquad (2.25)$$

and from

$$|(\mathbf{k}_{\rm i} + \mathbf{G})_{\perp}| = \sqrt{\frac{2m}{\hbar^2}} \left(\frac{\hbar^2 k_{\rm f}^2}{2m_0} + \phi - V_0\right),\tag{2.26}$$

 $k_{\rm i}$  and then the dispersion can be obtained.

### 2.2.2 Global band structure calculation: empirical pseudo potential method

Time-independent part Schrödinger equaiton in a periodic potential  $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})(\mathbf{R}$  is an arbitrary lattice vector) is written as

$$\mathscr{H}\psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}).$$
(2.27)

And we write a solution of the above in the Bloch form with dropping the band index as

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}). \tag{2.28}$$

The lattice periodic part  $V(\mathbf{r})$ ,  $u_{\mathbf{k}}(\mathbf{r})$  can be written in the Fourier expansion with the reciprocal lattice vector  $\mathbf{G}$  as

$$V(\boldsymbol{r}) = \sum_{\boldsymbol{G}} V_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}, \quad u_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} C_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}.$$
(2.29)

Substituting (2.28) and (2.29) into (2.27), we obtain

$$\sum_{\boldsymbol{G}} \left[ \left\{ \frac{\hbar^2}{2m} (\boldsymbol{k} + \boldsymbol{G})^2 - E) \right\} C_{\boldsymbol{G}} + \sum_{\boldsymbol{G}'} V_{\boldsymbol{G} - \boldsymbol{G}'} C_{\boldsymbol{G}'} \right] e^{i(\boldsymbol{k} + \boldsymbol{G}) \cdot \boldsymbol{r}} = 0.$$

Because each term in the summation of G should be zero, we obtain simultaneous equations for  $\{C_G\}$  as

$$\sum_{\mathbf{G}'} \left[ \left\{ \frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 - E \right\} \delta_{\mathbf{G}\mathbf{G}'} + V_{\mathbf{G} - \mathbf{G}'} \right] C_{\mathbf{G}'} = 0,$$
(2.30)

and the condition for the existence of non-trivial solution is

$$\left| \left[ \left\{ \frac{\hbar^2}{2m} (\boldsymbol{k} + \boldsymbol{G})^2 - E \right\} \delta_{\boldsymbol{G}\boldsymbol{G}'} + V_{\boldsymbol{G}-\boldsymbol{G}'} \right]_{\boldsymbol{G}\boldsymbol{G}'} \right| = 0.$$
(2.31)

If we can perform perfect expansion of (2.30) in an actual crystal, the solution of the secular equation (2.31) gives the accurate band structure  $E(\mathbf{k})$ . Equation (2.31) tells that requirements for this calculation are the coefficient  $V_G$  of Fourier expansion of periodic lattice potential.

In pseudo potential method we calculate "effective"  $V_G$  under the following concepts.

(1) Structures of valence band and conduction band below and above the Fermi level determine the properties of semiconductors. The outermost electrons of consisting atoms are forming these bands. The inner electrons are strongly bound around nuclei and can be included into the periodic crystal potential. Hence we apply the above secular equation only for the outermost electrons.

(2) (Characteristic for pseudo potential method) In the vicinity of nuclei,  $V(\underline{\mathbf{r}})$  can be approximated with r being the distance from the nucleons and Z being the atomic number as

$$V(r) = Ze/r,$$

and the outermost electron wavefunction should have stronger space modulation in the amplitude. On the other hand, far from the nucleus, the inner electrons (let the number be  $Z_c$ ) screens the potential and the effective atomic number decreases to  $Z' = Z - Z_c$ . Furthermore if we take into account the electron-electron mutual interaction in, *e.g.*, mean field approximation, the screening gives faster decay of potential than  $r^{-1}$  resulting in much weaker spatial modulation of the wavefunction. As we have seen in the tight-binding approximation (Sec.2.4), the band structure is dominated with the overlapping of wavefunctions in neighboring sites, that is, this weak potential part determines the band structure in practice.

If we perform the Fourier expansion of  $V(\mathbf{r})$  itself, strong spatial modulation around the nuclei introduces lots of high frequency components, which are nothing to do with the band structure. Such high frequency coefficients  $V_G$  not only introduce useless calculations but also make it difficult to solve the secular equation eq.(2.31).

The above consideration brings about the central concept of pseudo potential method. We look for a "pseudo potential" which simplifies the wave function around the nucleus but reproduces the tailing part of the wavefunction. Then obtain  $V_G$  for this pseudo potential and solve the secular equation (2.31).

The simplest example can be as in the left figure

Taking 
$$r_c$$
 to an appropriate value, we can approximately reproduce  
the tail of wavefunctions, keeping the eigen energies. Because the

(2.32)

 $W_n(r) = 0$   $(r < r_c), \quad W_p(r) = Z'e/r \quad (r \ge r_c).$ 

potential around the ion core is flat, a pseudo potential with small wavenumber expansion can be constituted. Summing up  $W_p$  on the positions of unit cells  $R_j$ , we obtain a pseudo potential for the crystal potential as

$$V_p(\mathbf{r}) = \sum_{j,\alpha} W_p^{\alpha}(\mathbf{r} - \mathbf{R}_j - \boldsymbol{\tau}_{\alpha}), \qquad (2.33)$$



**Fig.** 2.7 (a) Unit cell of a zinc blende crystal (GaAs). Alternate occupation of lattice points in diamond structure with group III (Ga) atoms and group IV (As) atoms. In the tetrahedron indicated by broken lines, apices are occupied with As and the center with Ga. (b) A primitive cell contains a Ga an an As atoms. Let the edge length of the unit cell in (a) a, and take the lattice point at the center between Ga and As then their coordinates are from the lattice point (a/8)(1, 1, 1) and (-a/8)(1, 1, 1) respectively.

where  $\alpha$  is the index of atomic positions in the unit cell and  $\tau_{\alpha}$  are relative positions of constituting atoms from a certain point in the unit cell.

Because potential (2.33) has the lattice periodicity, it can be Fourier expanded with wavenumber of reciprocal lattice points K.

$$v_{p}(\boldsymbol{K}) = \int \sum_{j,\alpha} W_{p}^{\alpha}(\boldsymbol{r} - \boldsymbol{R}_{j} - \boldsymbol{\tau}_{\alpha}) e^{-i\boldsymbol{K}\cdot\boldsymbol{r}} \frac{d\boldsymbol{r}}{V}$$
  
Let  $\boldsymbol{r}' \equiv \boldsymbol{r} - \boldsymbol{R}_{j} - \boldsymbol{\tau}_{\alpha}$ ,  $N$ : number of unit cells,  $\Omega$ : unit cell volume, From  $e^{-i\boldsymbol{K}\cdot\boldsymbol{R}_{j}} = 1$   
 $= \frac{1}{N} \sum_{j} e^{-i\boldsymbol{K}\cdot\boldsymbol{R}_{j}} \sum_{\alpha} e^{-i\boldsymbol{K}\cdot\boldsymbol{\tau}_{\alpha}} \frac{1}{\Omega} \int_{\Omega} W_{p}^{\alpha}(\boldsymbol{r}') e^{-i\boldsymbol{K}\cdot\boldsymbol{r}'} d\boldsymbol{r} = \sum_{\alpha} e^{-i\boldsymbol{K}\cdot\boldsymbol{\tau}_{\alpha}} \frac{1}{\Omega} \int_{\Omega} W_{p}^{\alpha}(\boldsymbol{r}') e^{-i\boldsymbol{K}\cdot\boldsymbol{r}'} d\boldsymbol{r}',$   
 $= \sum_{\alpha} e^{-i\boldsymbol{K}\cdot\boldsymbol{\tau}_{\alpha}} w_{p}^{\alpha}(\boldsymbol{K}).$ 
(2.34)

 $w_p^{\alpha}(\mathbf{K})$  is the Fourier transform of (2.32), and depends on the atomic species  $\alpha$ , that is the strength and functional form of nuclear potential, not on the crystal structure, and called **form factor**. On the other hand  $e^{-i\mathbf{K}\cdot\boldsymbol{\tau}_{\alpha}}$  depends only on the crystal structure and called **structure factor**. This separation of factors make it possible to estimate band structure based on the analogies between the crystals.

In the case of zinc blende structure, from Fig. 2.7(b), we can write  $\tau_1 = -a(1/8, 1/8, 1/8) = -\tau_2 \equiv \tau$ . Then (2.34) is written as

$$v_p(\mathbf{K}) = e^{i\mathbf{K}\cdot\boldsymbol{\tau}_1}v_p^1(\mathbf{K}) + e^{-i\mathbf{K}\cdot\boldsymbol{\tau}_1}v_p^2(\mathbf{K}) = (v_p^1 + v_p^2)\cos\mathbf{K}\cdot\boldsymbol{\tau} + (v_p^1 - v_p^2)\sin\mathbf{K}\cdot\boldsymbol{\tau}$$
$$= v_p^s(\mathbf{K})\cos\mathbf{K}\cdot\boldsymbol{\tau} + v_p^a(\mathbf{K})\sin\mathbf{K}\cdot\boldsymbol{\tau}.$$
(2.35)

Here  $v_p^s$  and  $v_p^a$  are symmetric and anti-symmetric part of the form factor  $v_p$ , sin and cosine functional part is the structure factor. In the case of diamond structure (like Si or Ge), from the symmetry  $v_p^1 = v_p^2$  and  $v_p^a = 0$ .

To obtain the value of form factor we need detailed functional form of the pseudo potential. Here comes the idea of "empirical" pseudo potential method, in which we rather *determine the form factors from experiments* as fitting parameters than to calculate them deductively from the specific form of pseudo potential.

#### 2.2.2.1 Global band structures of semiconductors with diamond and zinc blende structures

In the case of diamond structure  $v_p^a(\mathbf{K}) = 0$ , and because of the original concept of pseudo potential, we only need to calculate  $v_p^s(\mathbf{K})$  for reciprocal lattice points with small  $|\mathbf{K}|$ . Here we restrict  $|\mathbf{K}| \le \sqrt{11}$ , then  $a\mathbf{K}/2\pi = (000)$ , (111), (200), (220), (310) and their reversed points, 51 in total as in page ?? (for a while we drop the commas between the vector components. Accordingly the size of the matrix in the left hand side in (2.30) is  $51 \times 51$ .

The potential for  $|\mathbf{K}| = 0$  just shifts the energy and is set to zero. And from the above approximation, the terms for  $|\mathbf{K}| > \sqrt{11}$  are also dropped. The atomic potential is as in (2.32), supposed to have spherical symmetry. The form factors, which are the Fourier transform fo the potential, should also be a function of the absolute value of wavenumber. We thus

	$v_p^s(111)$	$v_{p}^{s}(220)$	$v_{p}^{s}(311)$	$v_p^a(111)$	$v_{p}^{a}(200)$	$v_{p}^{a}(311)$
Si	-2.856	0.544	1.088	0	0	0
Ge	-3.128	0.136	0.816	0	0	0
GaAs	-3.128	0.136	0.816	0.952	0.68	0.136
CdTe	-2.72	0	0.544	2.04	1.224	0.544

**Tab.** 2.3 Empirically obtained form factors of pseudo potential from optical reflection coefficients for representative fcc semiconductors in unit of eV. The values are taken from M L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).



**Fig.** 2.8 Band structures (solid lines )of (a) Si and (b) GaAs, obtained from the form factors in Tab.2.3. Broken lines show the results of linear muffin-tin orbital (LMTO) method, which is one of the simplest "first principles calculations". The two results are shifted to have the same value at the top of valence band. Energy zero E(k) = 0 is taken to the Fermi energy in the pseudo potential calculations.

only need to know the form factors for (111), (200), (220), (311) below the distance  $\sqrt{11}$ . Even among them for (200) the structure factor  $\cos \mathbf{K} \cdot \boldsymbol{\tau}$  is zero and we do not need to know the value or can put it as zero. And the last three are required. Then form factors are determined to fit to experimentally measured quantities. A way for further decreasing of the number of parameters is to determine  $r_c$  in (2.32) to obtain  $v_p^s(\mathbf{K})$  and perform the iteration to explain the experiments. Table 2.3 shows three parameters  $v_p^s(\mathbf{K})$  chosen as to fit to the optical reflection coefficients in experiments for representative diamond and zinc blende structure semiconductors.

Procedures of pseudo potential calculation for zinc blende semiconductors are similar to the above though  $v_p^a(\mathbf{K})$  is now finite. In the case of GaAs in Tab.2.3, since Ga and As locate in the both sides of Ge in the periodic table, the value of Ge is also adopted for  $v_p^s(\mathbf{K})$ . From (2.35), the anti-symmetric term is proportional to  $\sin \mathbf{K} \cdot \boldsymbol{\tau}$ , there is thus no contribution from (220), and those from (110), (200), (311) should be considered. Table 2.3 shows the results to reproduce optical measurements. Similarly in the case of a II-VI semiconductor CdTe, the value of Sn (gray tin) for  $v_p^s(\mathbf{K})$  is adopted and others are obtained from experiments.

Having the values of  $v_p(\mathbf{K})$ , we substitute them into (2.31) and solve the eigenvalue problem of the 51×51 matrix and obtain  $E(\mathbf{k})$ . Global band structures thus obtained are shown in Fig. 2.8.

Because the above calculations do not take care of the spin-orbit interaction, which actually has important contributions





to the band structure, in particular at the top of valence band, the three bands are degenerated. In pseudo potential calculations which take into account the spin-orbit contribution, one of the three branches goes down. In the results for Ge shown in Fig. 2.9, clear spin-orbit splitting at the top of the valence band is observed.

Above obtained band structures of Si, GaAs, Ge (representative diamond, zinc blende type semiconductors) are shown in Fig. 2.8 and Fig. 2.9. The bottoms of conduction band of Si are close to X-points though a bit inside the first Brillouin zone, while it is at  $\Gamma$ -point in GaAs and they are at L-points in Ge. Schematic drawing of equipotential surfaces are thus shown in Fig. 2.10, which are probably familiar to the readers. From Fig. 2.8 and Fig. 2.9, we see that the expressions in Fig. 2.10 are a bit exaggerated.

As can be seen above, Si has six equivalent bottoms in the first Brillouin zone, which are called **valleys**. In metallic doped n-type samples, the number of Fermi surfaces is that of valleys, to which we should pay attention in performing, *e.g.*, some integration over the Fermi surfaces. GaAs has a single valley at *Gamma*-point and the effective mass is almost isotropic. Ge has valleys at *L*-points and there are 8 equivalent *L*-points just at the Brillouin zone boundaries. Hence each valley is divided by the neighboring zones and the effective valley number is 4.

### 2.2.3 Band structure at band edges: Effective mass

When a Bloch type electron wavefunction  $\psi_{nk}(\mathbf{r})$  has a dispersion relation  $E_n(\mathbf{k})$ , the group velocity is written as

$$\boldsymbol{v}_n(\boldsymbol{k}) = \hbar^{-1} \nabla_{\boldsymbol{k}} E_n(\boldsymbol{k}). \tag{2.36}$$

Hence, the acceleration is given as

$$\frac{d\boldsymbol{v}_n}{dt} = \frac{d\boldsymbol{k}}{\hbar dt} \cdot \nabla_{\boldsymbol{k}} (\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k})) = \frac{\nabla_{\boldsymbol{k}}}{\hbar^2} \sum_{j=x,y,z} \frac{\partial E_n(\boldsymbol{k})}{\partial k_j} F_j.$$
(2.37)

Here,  $F = dp/dt = \hbar dk/dt$  is a vector of a "force". Now we define the **inverse effective mass tensor**  $1/m^*$ , which is the inverse matrix of the **effective mass tensor** with

$$\left(\frac{1}{m^*}\right)_{ij} \equiv \frac{1}{\hbar^2} \frac{\partial^2 E(\mathbf{k})}{\partial k_i \partial k_j}.$$
(2.38)

Then (2.37) can be re-written as

$$\frac{dv_i(\mathbf{k})}{dt} = \sum_j \left(\frac{1}{m^*}\right)_{ij} F_j,\tag{2.39}$$



**Fig.** 2.10 Schematic drawings of equipotential surfaces for (a) Si, (b) GaAs, (c) Ge conduction band valleys from the information obtained in the pseudo potential calculation of Fig. 2.9. In the case of Ge, which has the valleys at L-points, because the region is limited to the first Brillouin zone, the boundaries cut the centers of the spheroidal valleys.

which is equivalent to

$$F_i = \sum_j m_{ij}^* \frac{dv_j(\boldsymbol{k})}{dt}.$$
(2.40)

For simplicity, we consider an energy band with an isotropic energy dispersion  $E(k) = ak^2$ .  $m^*$ , in general is a tensor, becomes a scalar  $\hbar^2/(\partial^2 E(k)/\partial k^2) = \hbar^2/2a$ . Let it be more specific. Consider the case of eq. (1.9), where a gap opens up in the nearly free electron approximation (NFEA). Around  $\Delta k \sim 0$ ,

$$E_{\pm} \approx \epsilon_z \pm V_0 \left[ 1 + \frac{\epsilon_z}{2V_0} \left( \frac{\Delta k}{k_g} \right)^2 \right], \quad k_g \equiv \frac{\sqrt{2m_0 V_0}}{\hbar}, \quad (2.41)$$

which reads to the effective mass of

$$m^* = \pm \frac{\hbar^2}{2} \frac{2V_0}{\epsilon_z} \frac{2m_0}{\hbar^2} = \pm \frac{2V_0}{\epsilon_z} m_0 = \pm \frac{\epsilon_g}{\epsilon_z} m_0.$$
(2.42)

Here  $\epsilon_z$  is the band width,  $\epsilon_g$  the band gap. In this naive approximation, the ratio between the band width and the band gap determines the effective mass, namely, the wider the energy gap in comparison with the band width, the heavier the effective mass. This is a kind of "toy model" but can predict at least some trend in the effective masses in the same type of energy bands, *i.e.*, withe the same symmetry at the same point in the reciprocal lattices. For example we can see such tendency in the effective masses at  $\Gamma$  point of conduction band in GaAs, InP, InAs.

### 2.2.4 Measurement of band-edge structure: cyclotron resonance

The information given from ARPES introduced in Sec. 2.2.1 is limited to the region below  $E_F$  as we can see from the principle. And the precision is, at present, not enough for quatitative discussion on electric conduction or optical actions. Cyclotron resonance has long been used as a means of experimentally obtaining information on the band-edge of the conduction band and valence band. This method is based on the fact that the motion of the particle with charge q and mass m in the magnetic field with flux density B projected to the plane perpendicular to the magnetic field is a circular motion with cyclotron frequency

$$\omega_{\rm c} = \frac{qB}{m}.\tag{2.43}$$

Let's consider in classical approximation. The motion of equation for the particle with charge q and effective mass tensor  $\overleftarrow{m}$  is

$$\overleftarrow{m}\frac{d\boldsymbol{v}}{dt} + \frac{\overleftarrow{m}\boldsymbol{v}}{\tau} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}).$$
(2.44)

The electric field of microwave oscillates as  $Ee^{-i\omega t}$  and the velocity as  $ve^{-i\omega t}$  then

$$\left(-i\omega + \frac{1}{\tau}\right) \overleftrightarrow{m} \boldsymbol{v} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}).$$
(2.45)

Once we assume the oscillatio in v as above, the resonance condition does not depend on E and we put E = 0 for simplicity. For simpler expression we write  $\omega' = \omega + i/\tau$ ,  $B = B(\alpha, \beta, \gamma)$ , and put

$$\overleftarrow{m}^{-1} = \begin{pmatrix} m_1^{-1} & 0 \\ 0 & m_2^{-1} \\ 0 & 0 & m_3^{-1} \end{pmatrix}.$$
(2.46)

When the lattice system is cubic or rhombic this is justified. Then the equation of motion is written as

$$i\omega'm_1v_x + q(v_yB_z - v_zB_y) = 0,$$
  

$$i\omega'm_2v_y + q(v_zB_x - v_xB_z) = 0,$$
  

$$i\omega'm_3v_z + q(v_xB_y - v_yB_x) = 0.$$

For the above to have non-trivial solution,

$$\begin{vmatrix} i\omega'm_1 & qB\gamma & -qB\beta \\ -qB\gamma & i\omega'm_2 & qB\alpha \\ qB\beta & -qB\alpha & i\omega'm_3 \end{vmatrix} = 0.$$
(2.47)

From the condition  $\omega_c \tau \gg 1$ ,

$$\omega_c = \frac{qB}{m_c} = qB\sqrt{\frac{m_1\alpha^2 + m_2\beta^3 + m_3\gamma^2}{m_1m_2m_3}}.$$
(2.48)

 $m_c$  is obtained from the experiment with eq.(2.43) and called **cyclotron mass**.

In experiment, the condition  $\omega_c \tau \gg 1$  should hold and for that the impurities in the specimen should be very small. At low temperatures the carrier concentration discussed later is very low, and carriers are excited by light illumination. In Fig. 2.11(a), we show an example of thus obtained cyclotron resonance in the absorption of 24 GHz microwave. Several peaks are observed and assigned to electron (excitation in conduction band) and to hole (excitation in valence band). While the electrons have negative charge, the holes have positive. Hence if we use circular polarized microwave, the absorption intensity changes according to the direction of rotation. The assignment was done as above.

In the practical analysis, the symmetry of the crystals tells that the constant energy surfaces should be spheroids. We thus take the spatial coordinate to the main axes of a spheroid. Then the effective mass is represented by a tensor with



diagonal terms: $m_l$  along the main axis and two  $m_t$  along the other two axes. Let  $\theta$  be the angle between the magnetic field and the main axis, (2.48) gives

$$\left(\frac{1}{m_c}\right)^2 = \frac{\cos^2\theta}{m_t^2} + \frac{\sin^2\theta}{m_t m_l}.$$
(2.49)

The magnetic field angular dependences of cyclotron mass are displayed for the conduction bands in Ge and in Si in Fig. 2.11(b) and (c) respectively. From the analysis, we can get the directions of spheroids (and the number from the crystal symmetry) and the values of  $m_t$  and  $m_l$ .

## 2.2.5 Band-edge structure calculation: k·p perturbation

k·p perturbation is an adequate method to obtain highly accurate band structures around band edges. Though in empirical pseudo potential method we can reproduce band structure from very few parameters with comparatively simple calculation, in k·p perturbation we need to increase the number of bands included in the calculation, which increases the dimension of matrices and large scale calculation is required.

The basics of k·p perturbation is eq.(1.4) in the first hour of this lecture. Substituting Bloch function  $e^{i\mathbf{k}\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})$  into the original Schrödinger equation, we obtain the equation for the lattice periodic part  $u_{n\mathbf{k}}(\mathbf{r})$ . In three dimensional space (1.4) can be written as

$$\left[-\frac{\hbar^2 \nabla^2}{2m_0} + V(\boldsymbol{r}) + \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} - i\frac{\hbar^2}{m_0} \boldsymbol{k} \cdot \boldsymbol{\nabla}\right] u_{n\boldsymbol{k}}(\boldsymbol{r}) = E_{n\boldsymbol{k}} u_{n\boldsymbol{k}}(\boldsymbol{r}).$$
(2.50)

Here the Bloch wavenumber  $\boldsymbol{k}$  is a parameter (c-number) and not an operator.

Now we redefine the unperturbed Hamiltonian  $\mathscr{H}_0$  and k-dependent eigen energy as

$$\mathscr{H}_0 \equiv -\frac{\hbar^2 \boldsymbol{\nabla}^2}{2m_0} + V(\boldsymbol{r}), \quad E'_n(\boldsymbol{k}) = E_{n\boldsymbol{k}} - \frac{\hbar^2 \boldsymbol{k}^2}{2m_0}$$

Then the perturbation term can be written as

$$\mathscr{H}'(\boldsymbol{k}) = -i\frac{\hbar^2}{m_0}\boldsymbol{k}\cdot\boldsymbol{\nabla} = \frac{\hbar}{m_0}\boldsymbol{k}\cdot\hat{\boldsymbol{p}},$$
(2.51)

from which we can easily guess the source of the naming " $k \cdot p$ ".

(2.51) is zero for  $\mathbf{k} = 0$ . Hence we set the unperturbed state as  $\mathbf{k} = 0$ . Assume that we obtain the exact eigenstates  $\{u_{j0}(\mathbf{r})\}$  for  $\mathbf{k} = 0$ , then they form a complete set and an eigenstate for finite  $\mathbf{k}$  can be expanded as

$$u_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{j=0}^{\infty} c_{nj}(\boldsymbol{k}) u_{j0}(\boldsymbol{r}).$$

These  $c_{nj}(\mathbf{k})$  can be obtained from the perturbation of  $\mathscr{H}'(\mathbf{k})$ . This is the concept of k·p perturbation.

In the above  $\mathbf{k} = 0$  is taken to the unperturbed point assuming some avoided level crossing due to the high symmetry. Such avoided level crossing results in  $\partial E/\partial k = 0$  and the approximation is practically good around the point because  $\mathcal{H}'(\mathbf{k})$  is small around it. Similar may occur in other points with high symmetry and k·p expansions around such points are also available. Also as we have seen so far, physical properties of semiconductors are determined with band structures around such symmetric points. We cannot perform, of course, the infinite summation hence cut the summation around the band n which is under consideration. The accuracy of the k·p perturbation usually determined by the number of bands taken into account.

#### (a) the case of non-degenerate $u_{i0}(r)$

$$u_{i\boldsymbol{k}}(\boldsymbol{r}) = u_{i0}(\boldsymbol{r}) + \sum_{j \neq i} \frac{\langle j | \mathscr{H}' | i \rangle}{E_i - E_j} u_{i0}(\boldsymbol{r}), \quad E_i(\boldsymbol{k}) = E_i(0) + \langle i | \mathscr{H}' | i \rangle + \sum_{j \neq i} \frac{|\langle i | \mathscr{H}' | j \rangle|^2}{E_i - E_j}$$
(2.52)

are obtained as the first order perturbation. Here we have used abbreviation  $|i\rangle$  for  $|u_{i0}(\mathbf{r})\rangle$ . From  $\langle i|j\rangle = \delta_{ij}$  and  $\langle i|\nabla|i\rangle = 0$ ,

$$E_i(\mathbf{k}) = E_i(0) + \frac{\hbar^2 \mathbf{k}^2}{2m_0} - \frac{\hbar^4}{m_0^2} \sum_{j \neq i} \frac{\langle i | \mathbf{k} \cdot \nabla | j \rangle \langle j | \mathbf{k} \cdot \nabla | i \rangle}{E_i - E_j}.$$
(2.53)

#### (b) the case $u_{i0}(r)$ has degeneracy

When  $u_{00}(\mathbf{r})$  has *n*-fold degeneracy, we take an orthogonal basis  $\{u_{00}^{j}(\mathbf{r})\}$   $(j = 1, \dots, n)$  and write the functions in short form as  $|0j\rangle$ . Perturbed wavefunction is approximated with the linear combination  $|u_{0\mathbf{k}}^{i}\rangle = \sum_{j=1}^{n} A_{ij}(\mathbf{k})|0j\rangle$ . Substituting this into (2.50) gives  $[\mathcal{H}_{0} + \mathcal{H}' - E_{0}(\mathbf{k})]u_{0\mathbf{k}} = 0$ . With taking inner product with  $|0i\rangle$ , equation

$$\sum_{j=1}^{n} A_{ij}(\mathbf{k}) [\langle 0i|\mathcal{H}_{0}|0j\rangle + \langle 0i|\mathcal{H}_{0}'|0j\rangle - \langle 0i|E_{0}(\mathbf{k})|0j\rangle]$$
  
= 
$$\sum_{j=1}^{n} A_{ij}(\mathbf{k}) [\langle 0i|\mathcal{H}'|0j\rangle + (E_{0} - E_{0}(\mathbf{k}))\delta_{ij}] = 0 \quad (2.54)$$

is obtained. The secular equation for this simultaneous equation to have non-trivial solution is

$$|\langle 0i|\mathcal{H}'|0j\rangle + (E_0 - E_0(\mathbf{k}))\delta_{ij}| = 0,$$
(2.55)

which gives the dispersion relation  $E_0(\mathbf{k})$ . From the solution  $A_{ij}(\mathbf{k})$ , we obtain approximate set of eigenfunctions corresponding to  $\mathbf{k}$ .

# 2.2.6 Spin-orbit interaction

For rigorous derivation of **spin-orbit interaction** we should go back to Dirac equation, for which we do not have enough time unfortunately. Here without any derivation, we adopt the Hamiltonian for spin-orbit interaction as

$$\mathscr{H}_{\rm so} = -\frac{\hbar}{4m_0^2 c^2} \boldsymbol{\sigma} \cdot \boldsymbol{p} \times (\nabla V). \tag{2.56}$$

And just discuss the effect on the band structure.  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  is a vector which has Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.57}$$

as its elements. (2.56) is added to (2.50) and with (1.5), we obtain

$$\begin{bmatrix} \frac{p^2}{2m_0} + V + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} \boldsymbol{k} \cdot \boldsymbol{\pi} + \frac{\hbar}{4m_0^2 c^2} \boldsymbol{p} \cdot \boldsymbol{\sigma} \times \nabla V \end{bmatrix} |n\boldsymbol{k}\rangle = E_n(\boldsymbol{k})|n\boldsymbol{k}\rangle,$$
$$\boldsymbol{\pi} \equiv \boldsymbol{p} + \frac{\hbar}{4mc^2} \boldsymbol{\sigma} \times \nabla V.$$
(2.58)

We expand the solution, again with the basis of the band bottom  $|\nu 0\rangle$ . This time we need to take care of spin freedom and write  $|\nu, \sigma\rangle \equiv |\nu 0\rangle \otimes |\sigma\rangle$ , and expand as  $|n\mathbf{k}\rangle = \sum_{\nu',\sigma'} c_{n,\nu\sigma} |\nu',\sigma'\rangle$ . With taking inner product with  $\langle \nu,\sigma|$ , we obtain the eigen equation as

$$\sum_{\nu',\sigma'} \left\{ \left[ E_{\nu'}(0) + \frac{\hbar^2 k^2}{2m} \right] \delta_{\nu\nu'} \delta_{\sigma\sigma'} + \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{P}_{\sigma\sigma'}^{\nu\nu'} + \Delta_{\sigma\sigma'}^{\nu\nu'} \right\} c_{n\nu'\sigma'} = E_n(\mathbf{k}) c_{n\nu\sigma}.$$
(2.59)

where

$$\boldsymbol{P}_{\sigma\sigma'}^{\nu\nu'} \equiv \langle \nu\sigma | \boldsymbol{\pi} | \nu'\sigma' \rangle, \quad \Delta_{\sigma\sigma'}^{\nu\nu'} \equiv \frac{\hbar^2}{4m^2c^2} \langle \nu\sigma | [\boldsymbol{p} \cdot \boldsymbol{\sigma} \times (\nabla V)] | \nu'\sigma' \rangle.$$
(2.60)

The dispersion relation again is obtained with solving the eigen value problem. In this stage, it is often a good approximation to drop the spin-orbit part. In such cases it can be written as  $\pi = p$ ,  $P_{\sigma\sigma'}^{\nu\nu'} = \delta_{\sigma\sigma'} P_{\nu\nu'}$ .

# 2.2.7 Wavefunctions at $\Gamma$ -point in fcc semiconductors

In empirical  $k \cdot p$  method, without detailed knowledge of wavefunctions, the parameters required for the band calculation are extracted theoretically and the values are obtained from experiments. Many of such parameters are zero around highly symmetrical points making the calculation easier. Hence the knowledges of spatial symmetries in crystal and in atomic orbitals are important. Though the theory of space group gives systematic discussion to this problem, again due to the time limitation, we restrict ourselves to the discussion around  $\Gamma$ -point in fcc semiconductors.

Bravais lattice is fcc for group IV semiconductors with diamond structure and group III-V semiconductors with zinc blende structure. Here we name them "DZB" semiconductors. As is guessed from the structure in Fig.5.10(b), in chemical bond theory the crystal formation can be understood along covalent bonding between neighboring  $sp^3$  hybrid orbitals. In the group III-V semiconductors, for each atom to form  $sp^3$  hybrid, it needs to be ionized. Hence the crystals are also formed with the ionic bonding. The most effective atomic orbitals on the band structures in these semiconductors are sand p. In DZB structure, there are two atoms per a single lattice point in the simplest fcc structure (Fig.5.10). In Fig.5.9, substituting  $2 \times 4 = 8$  into Z, we see that the energy gap opens around the degeneracy points in the distance around  $\sqrt{3}$ from  $\Gamma$  point.

We consider a function  $|S\rangle$ , which has the lattice translational symmetry though also has the same angular symmetry as *s* orbital in the vicinities of nuclei. For that, we first take a linear combination of atomic orbitals (LCAO) of *s* orbital  $|s\rangle$ 

$$|u_s\rangle = \sum_{i,\beta} a_{i\beta} |s_{i\beta}\rangle,$$

where *i* is the index of unit cells,  $\beta$  is the relative index of atoms in a unit cell (as is in the pseudo potential calculation). Though the above function satisfies the crystal translational symmetry, it is not a solution for the Schrödinger equation with lattice potential. Hence we assume that we can modify the form of  $|s\rangle$  to make the linear combined function a solution for the Schrödinger equation with keeping the rotational symmetry in  $|s\rangle$  characteristic to the *s*-orbital. We write thus obtained LCAO wavefunction as  $|S\rangle$ , which must satisfy

$$\mathscr{H}_0|S\rangle = \left[-\frac{\hbar^2 \nabla^2}{2m_0} + V(\mathbf{r})\right]|S\rangle = E_c|S\rangle.$$
(2.61)

In the same way we define  $|X\rangle$ ,  $|Y\rangle$ ,  $|Z\rangle$ , which have angular symmetries of  $p_x$ ,  $p_y$ ,  $p_z$  respectively around nuclei, translational symmetry at the same time.

At  $\Gamma$ -point, the bottom of conduction band is mostly made from s orbitals while the top p orbitals. Hence, though the approximation is rough, we assume the above defined functions satisfy the unperturbed ( $\mathbf{k} = 0$ ) equation

$$\mathscr{H}_{0}|\zeta\rangle = \left[-\frac{\hbar^{2}\boldsymbol{\nabla}^{2}}{2m_{0}} + V(\boldsymbol{r})\right]|\zeta\rangle = E_{b}|\zeta\rangle, \qquad (2.62)$$

where  $\zeta \in \{S, X, Y, Z\}$ ,  $E_b$  is  $E_c$  for  $\zeta = S$  and  $E_v$  for others. It may be a problem whether such functions as  $|S\rangle$ ,  $|X\rangle$ ,  $\cdots$  actually exist. The space group theory says we can adopt lattice periodic functions with the same angular symmetries as s or  $p_{\alpha}$  orbitals around the point at which parabola with bottoms at  $(\pm 1, \pm 1, \pm 1)$  degenerate in the empty lattice approximation. <sup>\*1</sup>.

For the convenience to take into account the spin-orbit interaction, we transform basis from  $|X\rangle$ ,  $|Y\rangle$ ,  $|Z\rangle$  to

$$|+\rangle \equiv (|X\rangle + i|Y\rangle)/\sqrt{2}, \ |0\rangle \equiv |Z\rangle, \ |-\rangle \equiv (|X\rangle - i|Y\rangle)/\sqrt{2},$$

which correspond to eigen functions of angular momentum  $|p_{+1}\rangle$ ,  $|p_0\rangle$ ,  $|p_{-1}\rangle$  respectively. With the direct product of these four basis functions for the orbital part and two for the spin part  $(\uparrow, \downarrow)$ , eight basis functions in total, roughest  $k \cdot p$  perturbation calculation, in which the orbital degeneracy and the spin-orbit interaction are taken into account can be performed.

The perturbation Hamiltonian to  $|n\mathbf{k}\rangle$  is taken as

$$\mathscr{H}' + \mathscr{H}_{\rm SO} = -i\frac{\hbar^2}{m_0}\boldsymbol{k}\cdot\boldsymbol{\nabla} - \frac{\hbar}{4m_0^2c^2}\boldsymbol{\sigma}\cdot(\boldsymbol{p}\times\nabla V), \qquad (2.63)$$

in which we have dropped higher order terms from (2.58) and put  $\pi = p$ . The matrix elements between  $|S\rangle, |X\rangle, \cdots$  are

$$P \equiv \frac{\hbar}{m_0} \langle S|p_x|X\rangle = \frac{\hbar}{m_0} \langle S|p_y|Y\rangle = \frac{\hbar}{m_0} \langle S|p_z|Z\rangle,$$
(2.64)

$$\Delta \equiv -\frac{3i\hbar}{4m_0^2 c^2} \langle X | [\nabla V \times \boldsymbol{p}]_y | Z \rangle = (\text{cyclic replacement of } xyz)$$
(2.65)

and their conjugate elements. Others are zero due to the symmetries around nuclei.

Hence non-zero matrix elements of  $\mathscr{H}'$  are

$$\langle S\alpha | \mathscr{H}' | 0\alpha' \rangle = Pk_z \delta_{\alpha\alpha'}, \ \langle S\alpha | \mathscr{H}' | \pm \alpha' \rangle = \frac{\mp P}{\sqrt{2}} (k_x \pm ik_y) \delta_{\alpha\alpha'}, \tag{2.66}$$

and their conjugate elements  $\mathscr{H}'_{ji} = (\mathscr{H}'_{ij})^*$ , where  $\alpha, \alpha'$  are spin coordinates. As for  $\mathscr{H}_{SO}$ ,

and others are zero. From (2.62), unperturbed Hamiltonian  $\mathcal{H}_0$  has

$$\langle S\alpha | \mathscr{H}_0 | S\alpha' \rangle = \delta_{\alpha\alpha'} E_c, \quad \langle \{+, 0, -\}\alpha | \mathscr{H}_0 | \{+, 0, -\}\alpha' \rangle = \delta_{\alpha\alpha'} E_v. \tag{2.68}$$

<sup>\*1</sup> e.g. see Inui, Tanabe, Onodera, "Applied group theory" (Shokabo, 1976) Chapter 11 (in Japanese).



**Fig.** 2.12 Band structure for diamond and zinc blende semiconductors calculated from the lowest order k·p perturbation with adopting only S and P orbitals. Spin-orbit splitting exists though the heavy hole mass is the same as that of the vacuum electron, that is, the hole mass is negative in this calculation.

From the above we obtain the secular equation and thus the energy eigen values  $E_n(\mathbf{k})$ .  $\mathscr{H}'$  is an 8×8 matrix in the present basis though if we fix the wavenumber vector to z direction, *i.e.*,  $\mathbf{k} = (0, 0, k)$ , it becomes

$$\begin{bmatrix} H_d & 0\\ 0 & H_d \end{bmatrix}$$

thus is broken down to 4×4 matrices and

$$H_d = \begin{bmatrix} E_c & 0 & kP & 0\\ 0 & E_v - \Delta/3 & \sqrt{2}\Delta/3 & 0\\ kP^* & \sqrt{2}\Delta/3 & E_v & 0\\ 0 & 0 & 0 & E_v + \Delta/3 \end{bmatrix}.$$
(2.69)

From this the secular equation to obtain the eigenvalue  $\lambda$  is obtained as

$$\lambda = E_v + \frac{\Delta}{3},$$
  
$$(\lambda - E_c) \left(\lambda - E_v + \frac{2\Delta}{3}\right) \left(\lambda - E_v - \frac{\Delta}{3}\right) - |P|^2 k^2 \left(\lambda - E_v + \frac{\Delta}{3}\right) = 0.$$

In the second equation we approximate that the term of  $|P|^2k^2$  is small then obtain the energies for the conduction band  $E_c(\mathbf{k})$ , and the valence band  $E_{vj}(\mathbf{k})$  as

$$E_c(\mathbf{k}) = E_c + \frac{\hbar^2 k^2}{2m} + \frac{|P|^2 k^2}{3} \left[ \frac{2}{E_g} + \frac{1}{E_g + \Delta} \right],$$
(2.70)

$$E_{v1}(\mathbf{k}) = E_v + \frac{\Delta}{3} + \frac{\hbar^2 k^2}{2m_0},$$
(2.71)

$$E_{v2}(\mathbf{k}) = E_v + \frac{\Delta}{3} + \frac{\hbar^2 k^2}{2m_0} - \frac{2|P|^2 k^2}{3E_g},$$
(2.72)

$$E_{v3}(\mathbf{k}) = E_v - \frac{2\Delta}{3} + \frac{\hbar^2 k^2}{2m_0} - \frac{|P|^2 k^2}{3(E_g + \Delta)}.$$
(2.73)

The band structure around k = 0 thus far obtained is displayed in Fig. 2.12. Small mass of the conduction band, two different masses at the top of the valence band, and lowered energy of spin split-off band due to the spin-orbit coupling in the valence band, which properties are well known from optical measurements, cyclotron measurements, etc, are reproduced qualitatively though in particular, the heavier valence band mass is that of the vacuum electron, that is, the

hole effective mass is predicted to be negative apparently different from the real band structure. This is, of course, due to the coarse approximation, which is to the first order perturbation based on the degenerate four bands. The accuracy is enhanced by enhancing the order of perturbation to second, and by taking the surrounding bands into account. At present front of calculation, due to algorithm developments, and enhancement in computational performance have made it possible to perform calculations including over 20 bands and results with high accuracy which can be even used at comparatively high k[4].

Another way to utilize the result of  $k \cdot p$  "empirically" is, as is in the pseudo potential method, to represent the results of second order  $k \cdot p$  perturbation with a small number of parameters (*e.g.* Luttinger parameters) and to determine them fitting to the experiments. In the case of valence band in diamond and zinc blende semiconductors, the energies can be expressed as

$$E_v(\mathbf{k}) = E_v + \frac{\Delta}{3} + Ak^2 \pm \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 + k_x^2)},$$
(2.74)

$$E_{vsp}(\boldsymbol{k}) = E_v - \frac{2\Delta}{3} + Ak^2, \qquad (2.75)$$

and A, B, C are obtained from, e.g. cyclotron resonance.

# 2.3 Band structure of graphene

One of the ways to form a two dimensional electron system is to utilize two-dimensional crystals (two-dimensional materials). Graphene is the representative two-dimensional material. Graphene provides a good expample for the application of tight-binding calculation and we would like to see how the things go in a practical (though simplest) example.

The crystal structure of single-layer graphene is show in Fig. 2.13(a), which is a simple honeycomb structure of carbon atoms. The diamond drawn in the figure is the unit cell and the primitive lattice vectors and the primitive reciprocal lattice vectors are written as

$$\boldsymbol{a}_1 = \begin{pmatrix} \sqrt{3}a/2\\a/2 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 0\\a \end{pmatrix}, \quad \boldsymbol{b}_1 = \begin{pmatrix} 4\pi/\sqrt{3}a\\0 \end{pmatrix}, \quad \boldsymbol{b}_2 = \begin{pmatrix} -2\pi/\sqrt{3}a\\2\pi/a \end{pmatrix}.$$
 (2.76)

Henceforce we calculate the electronic states of graphene under simplest approximation. Because the approximation is rough, quantitative comparison with experiments is difficult. However, the results help understanding properties of graphene, *e.g.* the Dirac points appear at the Fermi level in pure graphene. Carbon belongs to group-IV and the outmost electrons exist in the orbitals 2s,  $2p_x$ ,  $2p_y$ ,  $2p_z$ . It is easy to see that these orbitals form  $sp^2$ -hybrids and the electronic states separate to  $\sigma$ -electrons  $(sp^2)$  and  $\pi$ -electrons  $(p_z)$ .  $\sigma$ -electrons form the honeycome through covalent bonding and the energy bands lie at low energy region. Then the electronic states placed around the Fermi level are  $\pi$ -electrons. Hence we consider Schrödinger equation on  $\pi$ -electrons on the honeycomb lattice.

We write the equation as

$$\psi = \mathscr{H}\psi, \tag{2.77}$$

and as Fig. 2.13(a), we separate the lattice sites to A-sites and B-sites on different sub-lattices. We consider a kind of tight-binding approximation between the two-sites. That is

$$\psi = \zeta_{\rm A} \psi_{\rm A} + \zeta_{\rm B} \psi_{\rm B}, \tag{2.78}$$

$$\psi_{\mathbf{A}} = \sum_{j \in A} \exp(i\mathbf{k}\mathbf{r}_j)\phi(\mathbf{r} - \mathbf{r}_j), \qquad (2.79a)$$

$$\psi_{\rm B} = \sum_{j \in B} \exp(i \boldsymbol{k} \boldsymbol{r}_j) \phi(\boldsymbol{r} - \boldsymbol{r}_j), \qquad (2.79b)$$



**Fig.** 2.13 (a) Two dimensional cryatal structure of graphene. Carbon atoms form a honeycomb lattice. It can be also viewed as an overlap of two face-centered square lattices placed at A and B positions. (b) Reciprocal lattice of (a).  $b_1$ ,  $b_2$  are the primitive reciprocal lattice vector corresponding to  $a_1$ ,  $a_2$ . The centtral point of the first Brillouin zone is  $\Gamma$ -point and as other points with high symmetries, K-point and M-point are indicated in the figure.

where  $\phi(\mathbf{r})$  is atomic wavefunction of  $\pi$ -electrons,  $\mathbf{r}_j$  are the lattice points. Here we write the matrix elements of the Hamiltonian between the each sub-lattice wavefunctions as

$$H_{\rm AA} = \langle \psi_{\rm A} | \mathscr{H} | \psi_{\rm A} \rangle, \quad H_{\rm BB} = \langle \psi_{\rm B} | \mathscr{H} | \psi_{\rm B} \rangle, \quad H_{\rm AB} = H_{\rm BA}^* = \langle \psi_{\rm A} | \mathscr{H} | \psi_{\rm B} \rangle. \tag{2.80}$$

And the number of atoms in the system is 2N, that is

$$\langle \psi_{\rm A} | \psi_{\rm A} \rangle = \langle \psi_{\rm B} | \psi_{\rm B} \rangle = N.$$
 (2.81)

Let  $\langle \psi_A | \psi_B \rangle$  be zero. We substitute (2.78) to (2.77). The condition of have non-trivial ( $\zeta_A$ ,  $\zeta_B$ ) givies the cecular equation

$$\begin{vmatrix} H_{AA} - NE & H_{AB} \\ H_{BA} & H_{BB} - NE \end{vmatrix} = 0.$$
 (2.82)

Lastly

$$E = (2N)^{-1} \left( H_{AA} + H_{BB} \pm \sqrt{(H_{AA} - H_{BB})^2 + 4|H_{AB}|^2} \right) \equiv h_{AA} \pm |h_{AB}|,$$
(2.83)

where we have used  $H_{AA} = H_{BB}$ , which comes from the symmetry, and we use lower cases for the quantities per atom with being devided by  $(2N)^{-1}$ .



**Fig.** 2.14 Vectors indicating three directional couplings between nearest neighbor carbon atoms.





**Fig.** 2.15 Left: surface plot of eq.(2.87). The figure shows the appearance of Dirac points, where the vertices of energy corns crash at the K-point. Upper: Schematical drawing of a Dirac point.

$$H_{\rm AB} = \sum_{l \in {\rm A}, j \in {\rm B}} \exp\left[i\boldsymbol{k}(\boldsymbol{r}_j - \boldsymbol{r}_l)\right] \langle \phi(\boldsymbol{r} - \boldsymbol{r}_l) | \mathscr{H} | \phi(\boldsymbol{r} - \boldsymbol{r}_j) \rangle_{\boldsymbol{r}}.$$
(2.84)

We further approximate that the off-diagonal matrix elements of  $\mathcal{H}$  just exist between the nearest neighbor sites. For the calculation we take the atom indicated as A in Fig. 2.13(a) as the center atom. The vectors from A to the nearest neighbor atoms 1, 2, 3 are  $d_i$  (i = 1, 2, 3) respectively. As is apparent from the figure,

$$\boldsymbol{k} \cdot \boldsymbol{d}_1 = \frac{k_x a}{\sqrt{3}}, \quad \boldsymbol{k} \cdot \boldsymbol{d}_2 = \left(-\frac{k_x}{2\sqrt{3}} + \frac{k_y}{2}\right) a, \quad \boldsymbol{k} \cdot \boldsymbol{d}_3 = \left(-\frac{k_x}{2\sqrt{3}} - \frac{k_y}{2}\right) a, \quad (2.85)$$

where  $a = |\mathbf{a}_1| = |\mathbf{a}_2|$ . The terms  $\langle \phi(\mathbf{r} - \mathbf{r}_l) | \mathscr{H} | \phi(\mathbf{r} - \mathbf{r}_j) \rangle_{\mathbf{r}}$  should be equal due to the symmetry and we write it as  $\xi$ . Consequently the residual resonant integral from the crystal structure is the repetition of the above and

$$h_{\rm AB} = \left(\sum_{j=1}^{3} \exp(i\boldsymbol{k} \cdot \boldsymbol{d}_j)\right) \boldsymbol{\xi}.$$
(2.86)

Substituting eqs.(2.85), (2.86) into eq.(2.83), we get the following expressio for the energy.

$$E = h_{AA} \pm \xi \sqrt{1 + 4\cos\frac{\sqrt{3}k_x a}{2}\cos\frac{k_y a}{2} + 4\cos^2\frac{k_y a}{2}}.$$
(2.87)

The second term is the perturbation from the nearest neighbor resonant integral, which vanishes at K-point in the reciprocal space

$$(k_x, k_y) = \left(0, \pm \frac{4\pi}{3a}\right), \quad \left(\frac{2\pi}{\sqrt{3}a}, \pm \frac{2\pi}{3a}\right), \quad \left(-\frac{2\pi}{\sqrt{3}a}, \pm \frac{2\pi}{3a}\right).$$
 (2.88)

We write  $k_y = 4\pi/3a$  and around  $k_x = 0$  (one of the K-points), eq. (2.87) can be approximated as

$$E\left(k_x, \frac{4\pi}{3a}\right) \approx h_{\rm AA} + \frac{\sqrt{3}\xi a}{2} |k_x|.$$
(2.89)

Namely, at the K-point the upper band has a lower pointed shape. Because the same for the lower band and as a result, at the K-point, as shown in Fig. 2.15, the band structure called **Dirac point**, which has no energy gap, no effective mass, appears.

Equation (2.87) is for a very simplified model. Just like a cosine band appeared in the tight-binding model in onedimension, the model itself does not have realistic meaning. However the model tells that the reason why we have the Dirac points at K-points is that the exsittence of three equivalent resonant integrals in eq. (2.86). The inference holds for the band calculation with any level precision since it is based on the symmetry. That meas the K-points in real graphene are really Dirac points.

# Appendix 2A: Band structure calculation based on density function theory

There are many electrons in an actual substance, and the wave function that expresses that state has fermion symmetry called antisymmetry with respect to the particle exchange operation. This causes an electron correlation effect. In addition, a Coulomb repulsive force acts between the electrons. In the semi-empirical band calculation, the effect of these electron-electron interactions is taken into account when the parameters of the one-electron band picture are obtained from the fit to the experimental values, but in the so-called first principle (ab initio) calculation, direct treatment of the electron-electron interaction is required. Calculations that incorporate electron-electron interactions require enormous amounts of calculation for high accuracy, computer resource savings are hence required. The density functional theory (DFT) is very advantageous in that point, thus used in many ab-intio calculations. Nowadays, calculation packages sometimes give us answers even without knowledge of calculations inside. Here, hoewever, we will have a brief look at very basics of ab-initio band calculations[6].

It has become clear that the ab-intio calculations with various approximations reproduce the qualitative features of the semiconductor band structure, but on the other hand, even the band gap, which is the most basic quantity cannot be reproduce without taking into account the quantum correlation effect though this is not an easy task. Incorporating the correlation effect properly is not easy even with DFT, and various theoretical modifications are made that can be said to be ad hoc, which is far from the "first principle" in some cases. It is necessary to pay attention to what kind of approximation is used for the calculation and how accurate it is.

### 2A.1 Kohn-Sham equation

Hohenberg and Kohn showed that the energy of interacting electron gas with electron density distribution  $\rho(\mathbf{r})$  in the external potential  $v(\mathbf{r})$  can be expressed with a universal functional of density  $F\{\rho(\mathbf{r})\}$  as

$$E\{\rho(\boldsymbol{r})\} = F\{\rho(\boldsymbol{r})\} + \int v(\boldsymbol{r})n(\boldsymbol{r})d\boldsymbol{r},$$
(2A.1)

and also showed that  $E\{\rho(\mathbf{r})\}$  takes minimum for the true electron density  $\rho(\mathbf{r})$ . The proof is for the case of ground state without degeneracy but the limitation was removed by Levy. Here we skip the proof.

Let the Hamiltonian without v(r) be

$$\mathscr{H}_{\mathrm{f}} = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \nabla_i^2 + \sum_{i>j}^{N} V(\boldsymbol{r}_i - \boldsymbol{r}_j), \qquad (2A.2)$$

and the density functional  $F\{\rho\}$  is generally written as

$$F\{\rho\} = \langle \Psi_N | \mathscr{H}_{\mathrm{f}} | \Psi_N \rangle. \tag{2A.3}$$

 $\Psi_N$  is the wavefunction to give  $\rho(\mathbf{r})$ . To obtain the form of  $F\{\rho\}$ , we write it in the form

$$F\{\rho\} = T\{\rho\} + U\{\rho\} + E_{\rm xc}^{(0)}\{\rho\}.$$
(2A.4)

 $T\{\rho\}$  is the kinetic energy,  $U\{\rho\}$  is the mean field expression of the Coulomb interaction between the electrons as

$$U\{\rho\} = \frac{e^2}{8\pi\epsilon_0} \int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}', \qquad (2A.5)$$

and is called Hartree term. The residucal  $E_{ex}\{\rho\}^{(0)}$  is called **exchange-correlation energy**.

Even the kinetic energy  $T\{\rho\}$  is difficult to be expressed explicitly with  $\rho$  and we make modification as follows. We consider an imaginary electron system without interaction in an effective potential  $v_{\text{eff}}(\mathbf{r})$ .

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\boldsymbol{r})\right]\psi_i = \epsilon_i\psi_i(\boldsymbol{r}).$$
(2A.6)

We also assume the electron density of this system coinsides with that of the interacting electron system.

$$\rho(\mathbf{r}) = \sum_{i} |\psi_i(\mathbf{r})|^2. \tag{2A.7}$$

The kinetic energy of the imaginary system is

$$T_{\rm S}\{\rho\} = -\frac{\hbar^2}{2m} \sum_{i=1}^N \int \psi_i^*(\boldsymbol{r}) \nabla^2 \psi_i(\boldsymbol{r}) d\boldsymbol{r}.$$
(2A.8)

We now renormalize  $E_{ex}$  to include the difference between  $T_{S}\{\rho\}$  and  $T\{\rho\}$  as

$$E_{\rm ex}\{\rho\} = E_{\rm ex}^{(0)} + T\{\rho\} - T_{\rm S}\{\rho\}.$$
(2A.9)

Then the total energy can be written as

$$E\{\rho\} = T_{\rm S}\{\rho\} + U\{\rho\} + E_{\rm ex}\{\rho\} + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r}.$$
(2A.10)

We minimize  $E\{\rho\}$  with the variation of  $\psi_i(\mathbf{r})^*$ . For that the normalization condition  $\langle \varphi_i | \varphi_i \rangle = 1$  leads to the introduction of Lagrange multipliers  $-\epsilon_i$ . Then the minimum condition is written as

$$\frac{\delta E\{\rho\}}{\delta \psi_i^*} = \left[ -\frac{\hbar^2}{2m} \nabla^2 + v(\boldsymbol{r}) + \frac{e^2}{4\pi\epsilon_0} \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' + \mu_{\text{ex}}(\boldsymbol{r}) - \epsilon_i \right] \psi_i(\boldsymbol{r}) = 0,$$
(2A.11)

where  $\mu_{ex}(r)$ , defined as

$$\mu_{\rm ex}(\boldsymbol{r}) = \frac{\delta E_{\rm ex}\{\rho\}}{\delta\rho(\boldsymbol{r})}$$
(2A.12)

is the quantity called exchange-correlation potential. The condition can be written in an eingenvalue equation as

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v(\boldsymbol{r}) + \frac{e^2}{4\pi\epsilon_0}\int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}d\boldsymbol{r}' + \mu_{\rm ex}(\boldsymbol{r})\right]\psi_i(\boldsymbol{r}) = \epsilon_i\psi_i(\boldsymbol{r}).$$
(2A.13)

Eq. (2A.13) is called Kohn-Sham equation.

In summary, the many-body effects are put into  $E_{ex}\{\rho\}$  in the above. The next step is how to calculate this term. In most frequently used method is **local density approximation** (LDA),

$$E_{\rm ex}\{\rho\} = \int \epsilon_{\rm ex}\{\rho(\boldsymbol{r})\}\rho(\boldsymbol{r})d\boldsymbol{r},$$
(2A.14)

then for  $\epsilon_{ex}\{\rho(\mathbf{r})\}\$ , the exchange-correlation energy for uniform electron gas with the density  $\rho$ . There are many other methods for the calculation.

# References

- [1] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. 98, 368 (1955).
- [2] M. L. Cohen and J. R. Chelikowsky, "Electronics Structure and Optical Properties of Semiconductors" (Springer, 1988).
- [3] C. Hamaguchi, "Basic Semiconductor Physics" 3rd ed. (Springer, 2017).
- [4] L. C. Lew Van Voon, M. Willatzen, "The k·p method" (Springer, 2009).
- [5] P. Hohenberg and W. Kohn, Phys. Rev. B 136, 864 (1964).
- [6] D. Sholl and J. A. Steckei, "Density Functional Theory: A Practical Introduction" (Wiley, 2009).