

# Physics of Semiconductors (1)

Shingo Katsumoto  
Institute for Solid State Physics, University of Tokyo

April 12, 2021

In the first half of this fiscal year (FY2021), I am assigned to the lecture on "Physics of Semiconductors." It's been a long time since I gave a lecture in one semester last time (8 years), and several new themes, which I want to introduce, have appeared. I am not very good at giving lectures like a machine-gun and want to take a course with a comparatively small amount of content. Though, maybe it's not enough for motivated students who want to learn a lot. Therefore, the notes follow the lecture and cover the advanced content that would open the eyes for more expansive fields. I would also like to introduce references for those who want to expand their studies.

## Chapter 1 General Properties of Semiconductors

### 1 What characterizes semiconductors?

Semiconductors refer to a form of solid that is usually classified by electrical conduction. Metals have large densities of states around the Fermi levels (that is, the Fermi surfaces exist) and are good conductors while insulators have their Fermi levels deep inside the wide energy gaps and interrupt electric currents. Semiconductors stand somewhere between them. They usually have comparatively narrow energy band-gaps, low but finite electric conductance at high temperatures, become insulating with lowering the temperatures.

However, such a viewpoint is not always useful nowadays. For example, a fine insulator with a large bandgap of 5.5 eV at room temperature, such as diamond, is also called a semiconductor, and devices are being made from it while materials with zero bandgap, such as graphene, are also important targets of researches in the semiconductor field. Rather, it seems to me that the property of "structure sensitive" fits better into the recent usage of the term "semiconductors." This is a long-used expression, which indicates the property that the electric conduction is sensitive to the ultra-small amount of impurities. Although, after the appearances of heterojunctions, MOS structures, superlattices, nanostructures, etc., I think the same expression is applicable to the sensitiveness of the transport properties on such real space structures.

In most cases, the object of such structure sensitive transportation is the electric charge, but recently the spin current in which the magnetic moment is transported by spin has also become an important research object. Research on spintronic devices is also active, and there is a possibility that some will eventually become practical.

For the spin current, which is the flow of magnetic momentum, the non-magnetic metals whose spins are canceled by the time-reversal symmetry, are like an empty space. It can be seen as a system similar to a semiconductor, which is a charge neutral space due to the charge cancellation of the nuclei and electrons. In fact, the inside of the metal is almost equipotential under normal experimental conditions, but a spin Hall spin current may exist. In spintronics, these systems are also "structurally sensitive" and look like semiconductors from the eye of semiconductor researchers. However, this is rather a unique view of myself, and usually, semiconductors are defined as those that are structurally sensitive in electric conduction.

### 2 Crystal Structures

#### 2.1 Lattice

A solid classified into crystal commonly has a spatially periodic structure of **basis**, which is also a certain structure of atoms. We represent such a state of matter as a **lattice**. "Spatially periodic structure" can be

represented as follows. An arbitrary point in a crystal with spatial coordinate  $\mathbf{r}$  has an infinite number of equivalent points  $\mathbf{r}'$ , which is represented, in the case of three-dimensional lattice, with three constant vectors  $\mathbf{a}_i$  ( $i = 1, 2, 3$ ) and three integers  $l_i$  ( $i = 1, 2, 3$ ) as

$$\mathbf{r}' = \mathbf{r} + \sum_{i=1,2,3} l_i \mathbf{a}_i = \mathbf{r} + \mathbf{R}. \quad (1.1)$$

Then the unit of the period is a certain set of atoms around  $\mathbf{r}$ . We take an arbitrary point in the unit. Such points form the lattice. We call such points as the lattice point.

The basis is the unit of the periodicity in the crystal and should be taken as to have the minimum number of atoms.  $\mathbf{a}_i$  in eq.(1.1) are called **primitive vectors**, while  $\mathbf{R}$  is called a **lattice vector**. The parallelepiped with  $\mathbf{a}_i$  as the edges contains a single basis is called **primitive cell**, with which we can fill up the space without gap. Primitive vectors often can be taken in multiple ways and usually taken as to make the symmetry of the lattice highest. A primitive cell is defined as a polygon which contains single basis and fills up the entire space without gap. Then there are infinite ways to define a primitive cell other than the above mentioned parallelepiped. When a block with multiple primitive cells is taken as the unit of period and the periodic structure has a higher symmetry, the block is more convenient for the unit. We then consider a **unit cell**, which may consist of single or multiple primitive cells.

As mentioned above, a crystal is composed of a unit structure and a lattice. The example of diamond structure, which often appears in group-IV semiconductors, is illustrated in Fig.1.1. In Fig.1.1(a) the atomic positions are indicated by middle-sized spheres, of which colors (black and white) indicate two different atomic sites in the crystal. The basis is composed of a black and a white atoms and a primitive cell can be taken as to contain these two sites. A point in the primitive cell, *e.g.* the position of black atom, can be taken as the lattice point. The consequent lattice is, as shown in (b), **face centered cubic** (fcc). Let  $\mathbf{e}_{x,y,z}$  be the unit vectors of the Cartesian coordinate system, then the primitive vectors can be taken as

$$\mathbf{a}_1 = \frac{a_0}{2}(\mathbf{e}_x + \mathbf{e}_y), \quad \mathbf{a}_2 = \frac{a_0}{2}(\mathbf{e}_y + \mathbf{e}_z), \quad \mathbf{a}_3 = \frac{a_0}{2}(\mathbf{e}_z + \mathbf{e}_x). \quad (1.2)$$

The primitive cell of the parallelepiped spanned by the vectors in eq.(1.2) is drawn with solid lines in Fig.1.1(a). On the other hand, the cubic drawn in the figure is often taken as a unit cell.

The lattices are classified by seven **crystal system** and additional lattice point (no point, face-centered, body-centered, base-centered) into 14 species of **Bravais lattice**.

## 2.2 Bravais lattice

The number of crystal structures is huge, maybe infinite, if we count, *e.g.* differences in molecular structures of organic crystals. On the other hand, the number of independent lattice structures is as small as 14 as shown in Fig.1.2. These 14 lattices are called three dimensional **Bravais lattice**.

The definition of the Bravais lattice classification is based on the discussion of spatial symmetry. The spatial symmetry of a manifold is defined by whether the manifold is invariant for the symmetry operations, such as rotation, reflection, translation, etc. For detailed discussion see, *e.g.* Ref.[1, 2]. Here we briefly summarize how we reach the 14 Bravais lattice.

We first classify the lattices with the relative lengths of primitive translational vectors  $a_1$ ,  $a_2$ ,  $a_3$  and the angles defined by two edges  $\theta_{12}$ ,  $\theta_{23}$ ,  $\theta_{31}$  (see the right-down inset of Fig.1.2). And for the rotational symmetry

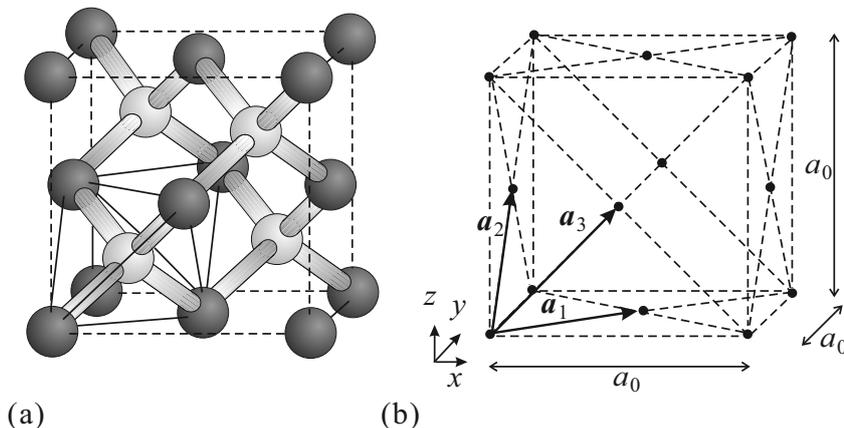


Figure 1.1: (a) Diamond structure. The circles correspond to atomic positions, while the thin cylinders correspond to covalent bonds. There is a single atom species though the two positions identified with colors, are different. The solid lines indicate the primitive cell. (b) Face-centered cubic lattice of the diamond crystal of (a).  $\mathbf{a}_{1-3}$  are the primitive vectors.

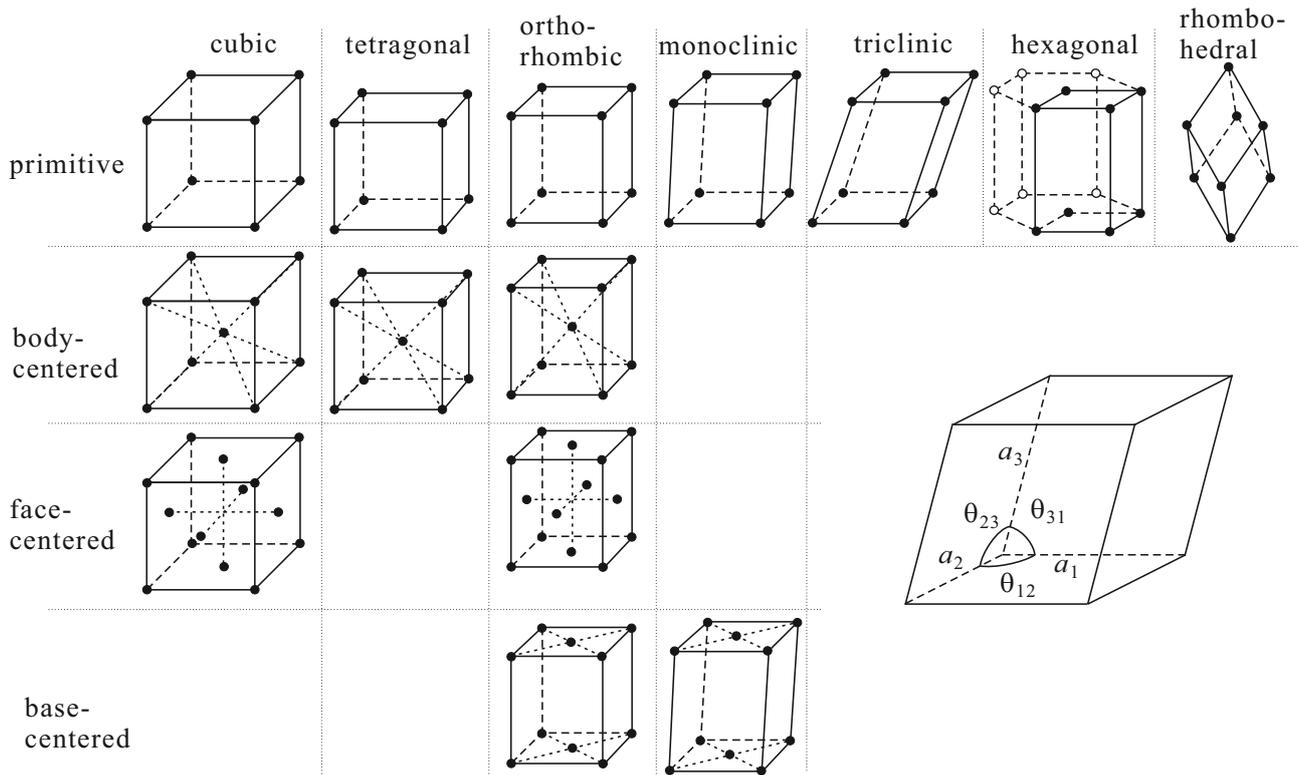


Figure 1.2: Bravais lattices of three-dimension. The parameters for the classification are illustrated in the right-down space.

around a primitive translational vector, the angle  $\pi/2$  is a special value and whether the angles  $\theta_{ij}$  are  $\pi/2$  or not is the other condition. These conditions leads us to the classification in Tab.1. This classification is called **crystal systems**. In three dimension, possible crystal systems are seven species in in Tab.1 and on the first column in Fig.1.2.

In the classification of crystal system, the focus is on the symmetry in the positions of neighboring lattice points. There are, however, some cases, in which we need to consider the relation between the next neighboring lattice points. For example, we take two (simple) cubic lattices and put them so as to lattice points of one of them are placed to the center of cubic in the other lattice. In this case if we look at the neighboring relation, the cubic symmetry seems to be lost but the second next ones are originally in cubic symmetry and the lattice is still classified into cubic crystal system but contains an additional lattice point at the center of cubic. The lattice, which contains an additional lattice point at the center of simple cubic, is called **body-centered cubic** (bcc).

In this way, the additional lattice points to the simple crystal system is another condition for the classification. The positions of such additional points are face-centered, body-centered and base-centered. As a consequence, we get 14 Bravais lattice shown in Fig.1.2.

Because we have ambiguity in taking “lattice”, the classification by Bravais lattice also has ambiguities. For a simplest example, in an fcc lattice crystal, if we take the unit as a single face-centered cubic, then the lattice

	$\theta_{12}$	$\theta_{23}$	$\theta_{31}$	$a_1, a_2, a_3$
cubic	$\pi/2$	$\pi/2$	$\pi/2$	$a_1 = a_2 = a_3$
tetragonal	$\pi/2$	$\pi/2$		$a_1 = a_2 \neq a_3$
orthorhombic	$\pi/2$			$a_1 \neq a_2 \neq a_3$
monoclinic	$\pi/2$	$\pi/2$		
triclinic	$\pi/2$			
hexagonal	$\pi/2$	$2\pi/3$		$a_1 = a_2$
rhombohedral(trigonal)	$\theta_0$	$\theta_0$	$\theta_0 \neq \pi/2$	$a_1 = a_2 = a_3$

Table 1: Conditions for Bravais lattice classification in three dimension. The definitions of the parameters are shown in the right-down panel in Fig.1.2.

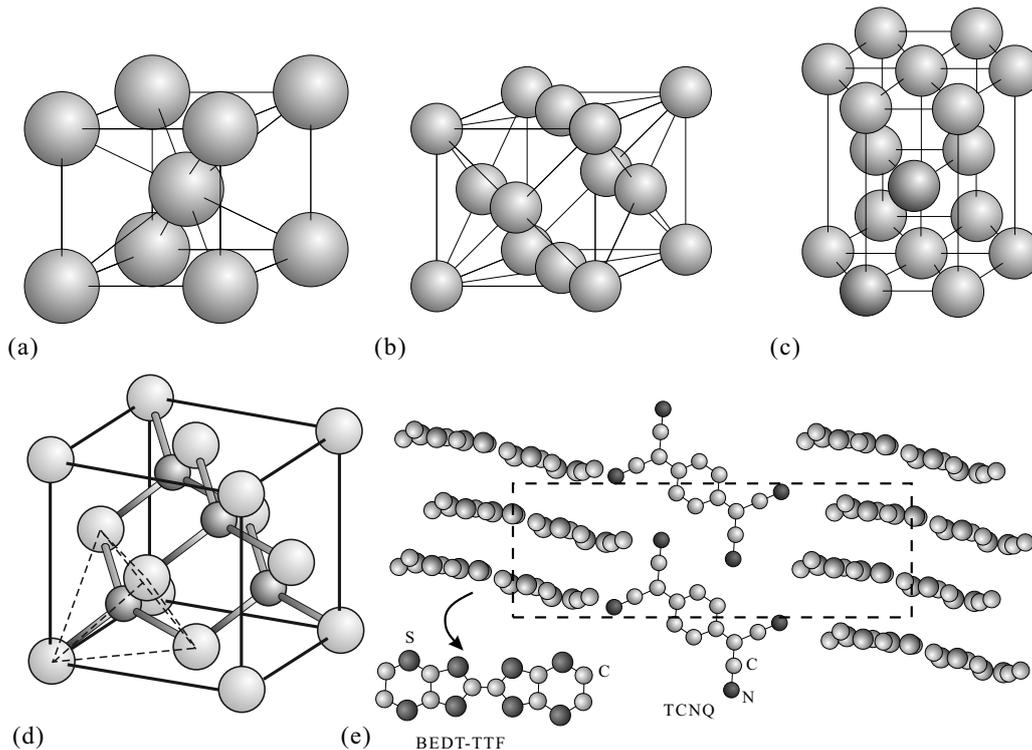


Figure 1.3: Examples of crystal structures. The centers of spheres indicate the positions of nuclei. (a) bcc-structure (Li, Na, Ba etc.), (b) fcc-structure (Al, Ni, Ag, Au etc.), (c) hcp-structure (Mg, Co, Cd etc.). These three often appear in crystals of metals. We get (a) and (b) with simply putting atoms at the lattice points of bcc and fcc lattices in Fig.1.2. (c) is classified into simple hexagonal lattice with taking two atomic positions which have a little darker color, as the basis. (d) is zinc-blende, which often appears in compound semiconductors. (e) is the structure of BEDT-TTF-TCNQ, which is one of organic semiconductors. It is difficult to see the structure in the BEDT-TTF molecules in the main panel. The inset at the left-down shows the molecule structure in the view from the direction vertical to the molecular plane.

is simple cubic. Also, the rhombohedral lattice can be viewed as a composite of three regular hexagonal prisms with  $120^\circ$  rotations to each other.

In Bravais lattices, the abbreviations fcc for face-centered cubic and bcc for body-centered cubic are frequently used. These lattices often appear in metal crystals and the primitive cells often consist of single atoms. Then for the crystal structures, fcc and bcc are also often used. The hexagonal close-packed structure shown in Fig.1.3(c) is also used as a crystal structure that appears approximately well in metal crystals, and the abbreviation hcp is used. The abbreviations for crystal structure are bcc, fcc, and hcp, but there is no hcp “lattice” in the sense defined here. That is, in Fig.1.3, the three atomic positions in the middle of the structure are not equivalent to the peripheral atomic positions, and one atom cannot be taken as a unit structure<sup>1</sup>. The basic structure can be taken as a combination of one atomic position in the upper and lower surfaces of Fig.1.3(b) and one atomic position in the middle, and the Bravais lattice is a hexagonal lattice.

Bravais lattice is a classification that focuses on the symmetry of the lattice and is important in the discussion of symmetry, but the fact that the symmetry of the lattice and the symmetry of the crystal are not the same means. It is clear from the fact that the unit cell is regarded as a lattice “point” in the lattice and the details in the unit cell are discarded. Let us take the diamond structure in Fig.1.1 again as an example. The position of the lower left apex of the regular tetrahedron, which is a part of the primitive cell, is the basic cell position, and the spatial arrangement is fcc in Fig.1.2. The gray colored parallelepiped is a primitive cell containing two atoms. In Fig.1.1, the difference of the two atomic positions is indicated by shades of colors. In the diamond structure, the atom species is the same for the two. In the case these are alternatively occupied by different species of atoms, *e.g.* Ga and As, the crystal structure is called zinc-blende (Fig.1.3(d)). That is, zinc-blende structure also belongs to fcc Bravais lattice. On the other hand, two atoms in the primitive cell are of the same

<sup>1</sup>On the web, many “hexagonal close-packed lattices” are searched, but in these explanations, the term “lattice structure” was used for “crystal structure”, and this combination was created. Also, “closest packing” mathematically means packing the spheres most densely. Crystals that have a mathematically perfect hcp structure are not known in real atoms because of their anisotropy.

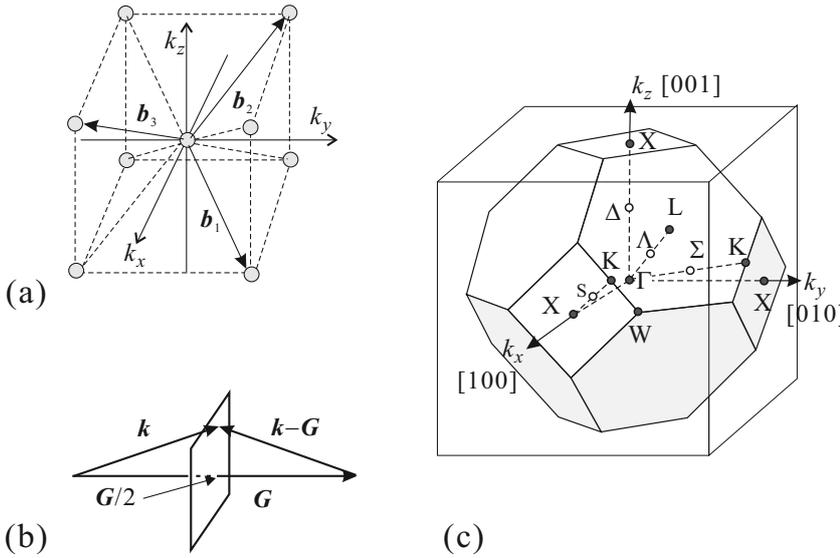


Figure 1.4: (a) Reciprocal lattice of fcc-lattice. The structure is classified to bcc. (b) A way to “cut” the reciprocal space to obtain Brillouin zone. Namely, cut at  $\mathbf{G}/2$  with the plane perpendicular to  $\mathbf{G}$ , where  $\mathbf{G}$  is a reciprocal lattice vector. (c) The first Brillouin zone obtained with the cuttings described in (b). The points indicated as  $\Gamma, X, L, \dots$  are the points with high symmetries.

species in the diamond structure, though of the different species in the zinc-blende structure. The former is symmetric for the inversion operation at the midpoint along the axis connecting the two positions while the latter is asymmetric.

Another example is in Fig.1.3(e), which shows an organic molecular crystal called (BEDT-TTF)<sub>2</sub>TCNQ. The atomic positions take a complicated form though the basis is single molecule and the lattice is triclinic. It is easy to understand the basis has strong anisotropy due to the atomic structure of the molecule and the symmetry of the crystal and that of the lattice is different. The symmetries of crystals are classified by the symmetry operations to 230 space groups.

### 2.3 Reciprocal lattice, Brillouin zone

Because the lattices of crystals have discrete translational symmetry, a potential  $U(\mathbf{r})$  ( $\mathbf{r}$  is spatial coordinate) caused by the lattice can be expanded in the Fourier series as

$$U(\mathbf{r}) = \sum_{\mathbf{G}} U_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}. \quad (1.3)$$

From the periodicity in (1.1),  $U(\mathbf{r} + \mathbf{R}) = U(\mathbf{r})$ . Then we obtain the condition for  $\mathbf{G}$  as

$$\mathbf{G} \cdot \mathbf{R} = 2\pi n \quad (n : \text{integer}), \quad \therefore e^{i\mathbf{G}\cdot\mathbf{R}} = 1. \quad (1.4)$$

The vectors  $\mathbf{G}$  which fulfill the condition (1.4) are called **reciprocal lattice vector**. Just like the real-space lattice, if we define **primitive reciprocal lattice vectors** as

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij} \quad (i, j = 1, 2, 3), \quad (1.5)$$

then we can write down  $\mathbf{b}_j$  ( $j = 1, 2, 3$ ) with  $|A| \equiv \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$  as

$$\mathbf{b}_1 = \frac{2\pi\mathbf{a}_2 \times \mathbf{a}_3}{|A|}, \quad \mathbf{b}_2 = \frac{2\pi\mathbf{a}_3 \times \mathbf{a}_1}{|A|}, \quad \mathbf{b}_3 = \frac{2\pi\mathbf{a}_1 \times \mathbf{a}_2}{|A|}. \quad (1.6)$$

A reciprocal lattice vector can be represented as  $\mathbf{G} = \sum_{i=1,2,3} h_i \mathbf{b}_i$  ( $h_i : \text{integer}$ ). Generally a function with the periodicity of a lattice can be Fourier expanded with the reciprocal lattice. It is legitimate to say a lattice and the corresponding reciprocal lattice are in the relation of Fourier transformation<sup>2</sup>.

As we considered primitive cells in spatial lattices, we can define the units of periodic repetition in reciprocal spaces. That is the **Brillouin zone**. A general way to obtain Brillouin zones is described in Fig.1.4(b). Let us see how to obtain the first Brillouin zone around the origin. The procedure is simply to cut the reciprocal space with planes containing the points  $\mathbf{G}/2$  and perpendicular to  $\mathbf{G}$ , where  $\mathbf{G}$  are the reciprocal lattice vectors starting from the origin. The minimum space (polyhedron) around the origin surrounded by such surfaces is

<sup>2</sup>As an analytic expression, it is enough to remember that the Fourier transform of a regular series of  $\delta$ -functions is, again, a regular series of  $\delta$ -functions. This is a very general principle. An example is the optical frequency comb.

the first Brillouin zone. This way of “cutting at  $\mathbf{G}/2$ ” will have meaning in considering the band structure, which will be discussed in the next chapter.

The example of fcc-lattice is shown in Fig.1.4. First, the primitive reciprocal lattice vectors are obtained from eq.(1.6). Then we find the reciprocal lattice for the fcc-lattice is bcc-lattice as shown in Fig.1.4(a). Next we apply the method in (b). The reciprocal lattice vectors pointing the nearest neighbor reciprocal lattice points are the primitive reciprocal lattice vectors  $\pm\mathbf{b}_1, \pm\mathbf{b}_2, \pm\mathbf{b}_3$ . There are equivalent eight planes which cut the vectors vertically at the midpoints. The polyhedron covered with these planes is a regular octahedron. However, the planes which cut the vectors pointing the next nearest neighbor reciprocal lattice points at the midpoints, also cut the octahedron around the vertices. The procedure thus results in the first Brillouin zone shown in Fig.1.4(c), where  $\Gamma, X, L, \dots$  indicate the points with high symmetry. The points are often used in the display of band structure.

## 2.4 Crystals often used as semiconductors

Which materials should be called “semiconductors” is a difficult problem, and some scholars propose the classification of “every material that is not metal”. In fact, diamond, which was a typical insulator a while ago, has recently completely established itself as a semiconductor. Here, let’s have a quick look at the simple and clear “crystals” of the spatial periodic structure, and those that are often used as semiconductors in the industry.

As specific examples, we take materials consist of comparatively small numbers of elements from Group II to Group VI in the periodic table. In the right table, we show the part of the periodic table under consideration with the electronic orbital occupation. We can guess from the table that the semiconductors composed of these elements takes similar lattice structures. Here we mainly introduce crystal structures.

II	III	IV	V	VI
${}^4\text{Be}$ $2s^2$	${}^5\text{B}$ $2s^2 2p$	${}^6\text{C}$ $2s^2 2p^2$	${}^7\text{N}$ $2s^2 2p^3$	${}^8\text{O}$ $2s^2 2p^4$
${}^{12}\text{Mg}$ $3s^2$	${}^{13}\text{Al}$ $3s^2 3p$	${}^{14}\text{Si}$ $3s^2 3p^2$	${}^{15}\text{P}$ $3s^2 3p^3$	${}^{16}\text{S}$ $3s^2 3p^4$
${}^{30}\text{Zn}$ $3d^{10}$ $4s^2$	${}^{31}\text{Ga}$ $3d^{10}$ $4s^2 4p$	${}^{32}\text{Ge}$ $3d^{10}$ $4s^2 4p^2$	${}^{33}\text{As}$ $3d^{10}$ $4s^2 4p^3$	${}^{34}\text{Se}$ $3d^{10}$ $4s^2 4p^4$
${}^{48}\text{Cd}$ $4d^{10}$ $5s^2$	${}^{49}\text{In}$ $4d^{10}$ $5s^2 5p$	${}^{50}\text{Sn}$ $4d^{10}$ $5s^2 5p^2$	${}^{51}\text{Sb}$ $4d^{10}$ $5s^2 5p^3$	${}^{52}\text{Te}$ $4d^{10}$ $5s^2 5p^4$

### 2.4.1 Group IV semiconductors

Elementary semiconductors of C, Si, Ge take **diamond structure** (Bravais lattice is fcc). The bonds in these crystals are dominated by covalent binding of  $sp^3$  hybrid orbitals. Silicon (Si) is of course the king of semiconductors in the industry. Tin (Sn) are metals in many phases but the form called  $\alpha$ -Sn (gray tin) is a semiconductor with diamond crystal structure.

Creation of low dimensional electron system is a big charm of semiconductor physics, and it is also very important in the semiconductor industry. In the case of Si, a metal-oxide-semiconductor (MOS) structure has long been used to create two-dimensional electron systems. Since the oxide layer generally takes an amorphous structure, the interfacial scattering probability of two-dimensional electrons is high, and it is difficult to obtain an electron system with high mobility. On the other hand, a two-dimensional electron system with high mobility has been realized by a method of forming a  $\text{p-n}$  heterojunction using mixed crystals of Si-Ge.

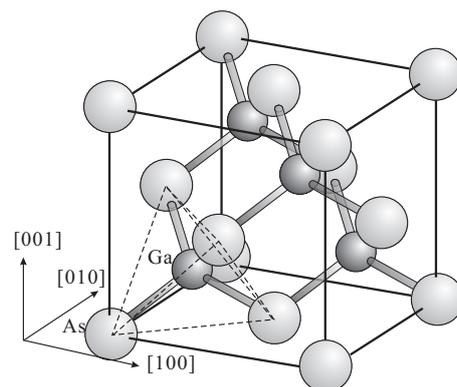
semiconductor	lattice constant $\text{\AA}$	energy gap (RT eV)	electron mass $m_0$	hole mass
C	3.56683	5.47	0.25	0.2
Ge	5.64613	0.66	1.64, 0.082	0.04, 0.28
Si	5.43102	1.12	0.98, 0.19	0.16, 0.49

### 2.4.2 III-V compound semiconductors

Semiconductors made by combining group III elements and group V elements on a one-to-one basis, group III and group IV atoms occupy the lattice points of the diamond structure alternately, that is, **zinc blende** structure. Al, Ga, In are often used as group III, and As, P, Sb, etc. are often used as group V. Many kinds of compound semiconductors are formed by these combinations, and more kinds of semiconductors can be synthesized by further mixing different kinds of elements to form mixed crystal. Strictly speaking, these are no longer crystals because they have lost the spatially regular structure, but most of the concepts in the crystals work well by considering some blunting due to lattice disorder.

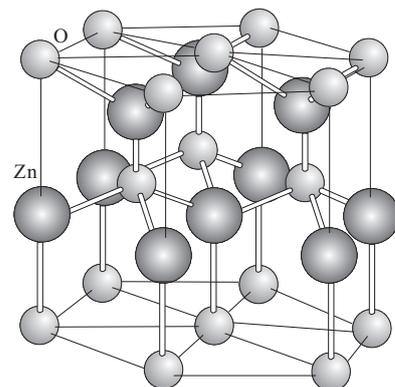
In reality, these compound semiconductors and mixed crystal semiconductors are often synthesized by epitaxial growth, and for this purpose, the crystal form and lattice constant should be similar between the heterogeneous semiconductors to be joined. This will be described later.

Many of III-V semiconductors have a direct gap at the origin of the reciprocal lattice space,  $\Gamma$  point, and are therefore often used for optical devices. In addition, there are many combinations that can form high-quality heterojunctions by epitaxial growth, and they are frequently used for devices for high-speed operation.



### 2.4.3 III-N compound semiconductors

In nitride semiconductors, whose applications have expanded rapidly for blue light emitting diodes, high-frequency, high-power devices, GaN, InN, and AlN are currently the main research targets. They take hexagonal Wurtzite crystal forms. They are usually grown by use of epitaxial growth and annealed at high temperature to improve their quality. The figure in the right shows the crystal structure of Wurtzite (for the case of ZnO).



### 2.4.4 II-VI compound semiconductors

Group II-VI semiconductors take various crystal forms such as zinc-blende, wurtzite, and chalcopyrite. There are various compounds such as ZnO and CdTe, and before GaN became the leading player in blue-color optical devices, the II-VI system was mainly studied as a candidate. ZnO is still considered to be a material that threatens the GaN system if the device characteristics and manufacturing method are improved because the material is easily available. The ZnO system tends to have a small structure such as nanotubes, and it is not easy to form it into a thin film device. On the other hand, its application as a nanostructure device is also attracting attentions. The Hg system is said to have a “negative bandgap” and has become well known for its use in constructing topological materials.

## 2.5 Organic semiconductor materials

Organic thin films as semiconductors are attracting attention because they are lightweight, flexible, and inexpensive. Most organic solids are molecular solids in which intermolecular bonds are formed by van der Waals forces. The qualities of organic semiconductors have been improved and the various concepts of semiconductor physics are now applicable to organic ones. However, it is often more realistic to consider the electronic state in the molecule and the solid state as an aggregate of them separately, reflecting the fact that they are molecular solids. Especially in the case of macromolecules, on one hand Bloch electrons and bands in the molecule are good approximation, on the other hand, the electrical conduction of the whole solid should be analyzed with the theory for amorphous solids developed in the 1970s and 1980s.

## 3 Crystal growth

In order to utilise the structural sensitivity of semiconductors as functions to explore condensed matter physics, to setup them as laboratories of quantum and many-body effects and to use them as devices, we need to obtain,

as the starting point, obtain crystals with very low concentrations of defects and impurities. For that, the original materials with ultra high purity, higher than those in reagents by orders, should be prepared with cheap cost, huge amount, in very short time, and with very low environmental load. The crystal growth is hence a high field in the semiconductor industrial science. The physics, the main issue of this lecture, is not directly connected to that field but I would like to introduce some in a very short time.

Crystal growth methods of inorganic semiconductors can be classified to one for three-dimensional bulk growth and another for two-dimensional growth on wafers of crystals cut from three dimensional ingot. The latter is called epitaxial growth.

### 3.1 Growth of bulk crystals

Mining and refinement of source materials are very important processes before the crystal growth, and we need to choose the best degree of material quality and refinement method considering the cost and the final product. In the case of crystalline silicon, it is said that astonishing purity of 11N(99.99999999%) is required for substrates of MOS-LSI<sup>3</sup>, which is called “semiconductor grade”.

On the other hand, a solar cell device generally has an area 10 orders of magnitude wider than that of MOS-LSI, and the tolerance for leakage current per area also differs by a few orders of magnitude. Hence for them, the purity of 6N~7N is enough under reduction of impurities that form non-radiative recombination centers or pn characteristics degrading deep levels. Such wafers are called in “solar grade”.

In the latter, usually low quality Si called “metal grade” is used as a starting material. There have been long term seekings for purification method with low power consumption and some new progress has been made though the world market is now dominated by companies which provides cheap wafers produced with traditional method in 2013. Such situation is largely affected by international affairs or economic atmosphere. I am sorry but must say that “basic researches” are affected by such political situation in reality.

Bulk single crystals of inorganic semiconductors are usually obtained from gradual solidification by cooling from high temperature melts. This is comparatively easy for single element semiconductor Si or Ge. In the growths of compound semiconductors, mixed melts of multiple elements should be prepared and the difference in melting point, vapour pressure and mutual solubility are the possible problems.

#### 3.1.1 Czochralski process

In **Czochralski (CZ)** process, as illustrated in Fig.1.5 a thin seed crystal is put down to the surface of a melt from source materials, and a thick cylindrical crystal is pulled up. The grown crystal is formed in a cylindrical shape because the seed is rotated during the pulling up growth process. This is a representative method to

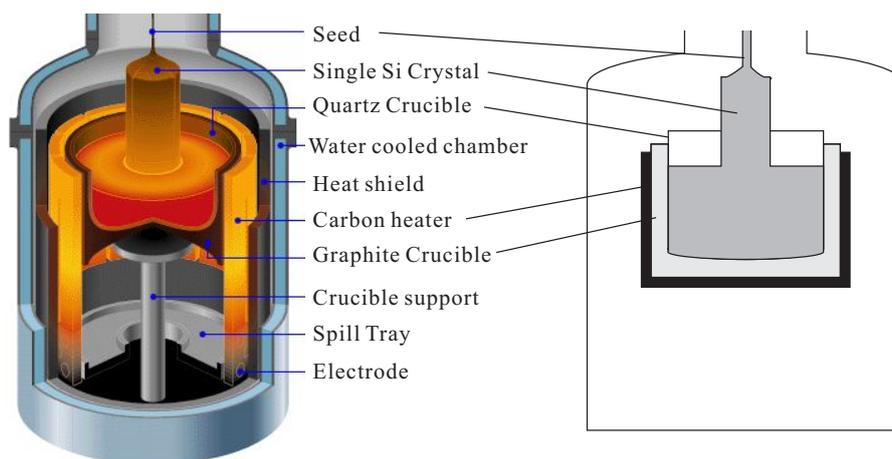


Figure 1.5: Schematic drawing of Czochralski process. Left: Three-dimensional schematic illustration. Right: Cross-sectional illustration.

From [http://people.seas.harvard.edu/~jones/es154/lectures/lecture\\_2/materials/materials.html](http://people.seas.harvard.edu/~jones/es154/lectures/lecture_2/materials/materials.html)

<sup>3</sup>Here they are using a special definition of “purity”. I have experienced that such an ultra-pure Si ingot contains a considerable amount of oxygen measured from low temperature magnetic susceptibility measurement with a SQUID magnetometer. 11N is hence the value on the ignorance of these impurities. Oxygen has little effect on logic LSIs but is a problem in the application for power devices.

obtain a dislocation free crystal of Si. The thin disk form popular for LSI wafer appears after slicing the cylindrical columnar shape.

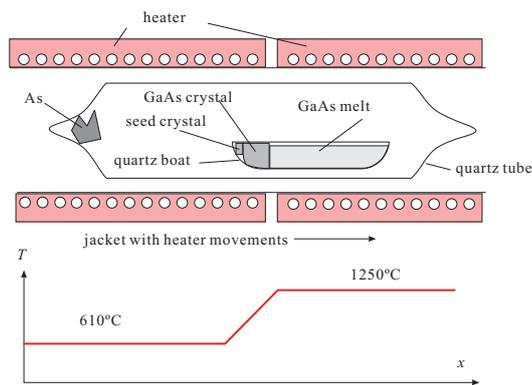


Figure 1.6: Schematic illustration of a boat method (horizontal Bridgeman process).

from one end the melt is frozen into a single crystal.

Figure 1.6 shows a schematic illustration of HB method for the case of GaAs. Initially a metallic solid As is set to one end of a quartz tube, which also contains a quartz made boat. In the beginning Ga melt and a seed crystal are in the boat. The side of the As metal is heated to  $610^\circ\text{C}$  while the other side to  $1250^\circ\text{C}$ . As sublimates severely above  $600^\circ\text{C}$  and gets into Ga melt forming GaAs melt. At  $1250^\circ\text{C}$ , GaAs is in melting phase and at  $610^\circ\text{C}$  in solid phase. As the furnace moves to the right in the figure, a GaAs single crystal is solidified from the end of the seed crystal to the right.

### 3.1.3 Zone melting method

As mentioned in the footnote in page E1-8, “ultrahigh purity” Si actually often contains high concentration of oxygen, which mainly comes from the crucibles during the growth. For power MOS FET or other devices in which such oxygen causes troubles, single crystals are grown by **floating zone** melting (FZ) method. In the initial stage, a rod of polycrystal with a high purity is prepared in standing manner and a seed crystal is put on top of the rod. At first a zone of the polycrystal rod from the top is melted e.g., by concentration of infrared beam with confocal method or by rf loss heating. The melt in contact with the seed crystal changes into single crystal and the melted zone slowly goes down to form a single crystal rod. During the process the melt does not touch any other materials and the high purity of polycrystal is kept. On the other hand, such big radiuses of grown rods as those in CZ method cannot be obtained.

## 3.2 Epitaxial growth of thin films

**Epitaxial growth**, in which thin crystal films are grown with deposition of materials onto crystal substrates, is classified into **liquid phase epitaxy** (LPE), **vapor phase epitaxy** (VPE), and epitaxy in vacuum or in very low pressure gas. Here I will pick up metal organic vapor phase epitaxy (MOVPE) and molecular beam epitaxy (MBE) from the number of epitaxial growth methods.

Such a primitive CZ method cannot generally be applied to compound semiconductors due to large difference in the evaporation pressure. Actually CZ method is often adopted in growth of III-V semiconductors GaAs, InP, GaP, etc. but not in the primitive form because the group V materials have much higher vapour pressures than those of the group III, resulting in the rapid escape of group V materials from the melts. Instead, **Liquid Encapsulated Czochralski** (LEC) process, in which the melt of the sources is encapsulated with an inert liquid like  $\text{B}_2\text{O}_3$ .

### 3.1.2 Boat method

Another popular method for bulk-growth of compound semiconductors is the one called “boat method”. The boat method is further classified into horizontal Bridgeman (HB) method and temperature gradient freeze method. In the former, a furnace with two temperature regions is moved along a boat, in which the source materials are melt, and

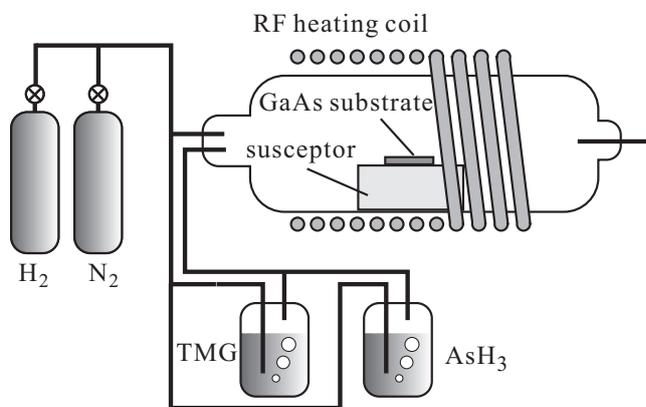
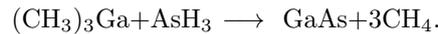


Figure 1.7: Extremely simplified schematic illustration of MOVPE (MOCVD) apparatus of GaAs deposition. The “susceptor” absorbs the power of RF and gets heat.

### 3.2.1 Metal-organic vapor phase epitaxy

Also often called as Metal organic chemical vapour deposition (MOCVD). Often used for the growth of compound semiconductors. Let us see the case of GaAs.

In general in epitaxial growth of a thin film crystal, component materials are carried onto the substrate with some carriers or with some other method, and react with the substrate surface to form single crystal films. Therefore the keys for the growth are the surface states of the substrate, the way of carrying the materials, the dynamics of deposited atoms, etc. In the case of MOVPE, the sources are carried by hydrogen and nitrogen gases. Ga is put on tri-methyl gallium ((CH<sub>3</sub>)<sub>3</sub>Ga, TMG), and As on arsine (AsH<sub>3</sub>). They are carried onto the substrate and decomposed into atoms on the surface by heating. Then they are chemically bonded to the surface atoms to form GaAs crystal films. Omitting all the intermediate chemical reactions and the initial and the final states can be written as



TMG and arsine have low vapor pressures and as shown in Fig.1.7, liquids of them are bubbled with hydrogen to be vaporized. Hydrogen gas is deoxidizing atmosphere for GaAs surface. Thus flat and high quality films can be grown though the actual chemical reaction is not so simple. Doping of impurities, growth of mixed crystals are possible with preparation of materials. All of metal organic gases of group II or III, arsine or phosphine of group V are explosive and at the same time nerve gases. They are extremely dangerous and should be treated with highest care and rigid safeguards.

### 3.2.2 Molecular beam epitaxy

Molecular beam epitaxy (MBE) is a representative growth method of ultra-thin semiconductor films. Characteristic features are: (1) deposition in ultra-high vacuum; (2) single crystalline substrates and various methods for surface cleaning; (3) heating of substrates to activate the motions of deposited atoms; (4) stoichiometric deposition of materials is not always required; (5) *in situ* characterization of grown films in various ways is possible because the growth front is always on the surface to the vacuum.

Figure 1.8(a) is a schematic illustration of an MBE machine, (b) shows a photograph of a real machine. In order for keeping ultra-high vacuum in the growth chamber, (a) “pre-evacuation chamber(s)” is used for loading and unloading of substrates. Molecular beam cells (Knudsen cells, K-cells; Langmuir cells, L-cells), which have sources of evaporation in them, are kept at intermediate temperatures for them to avoid adsorption of gas molecules. While the growth, the shrouds covering substrate, heating system and molecular

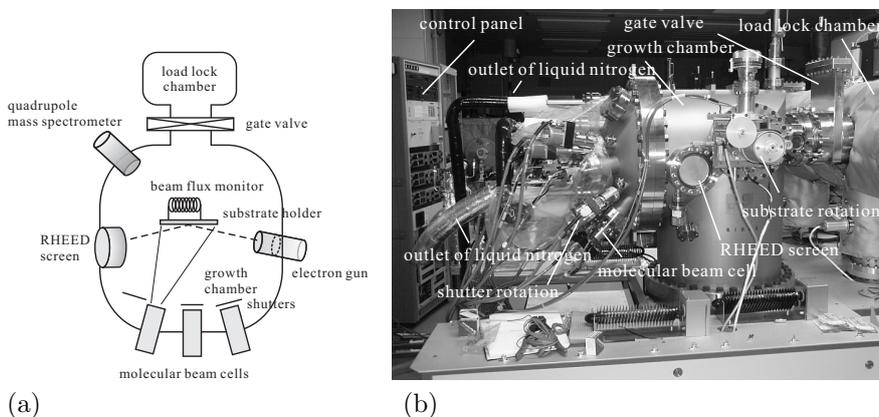


Figure 1.8: (a) Schematic illustration of an MBE machine. (b) Photograph of a real machine (RIBER S32).

beam cells are cooled down with continuous flow of liquid nitrogen to adsorb outgassed molecules. When evaporations are going on, the source molecules are inside the chamber and the total gas pressure increases, which makes the quality of vacuum obscure. To monitor the quality, we need a partial pressure gauge for gas species. That requirement, not 100 % is fulfilled by a mass spectrometer, with which partial pressure can be measured as a function of the ratio of the molecular mass to the charge.

Substrates for growths are introduced into the preparation (pre-evacuation) chamber after surface cleaning with chemical etching and protection of the cleaned surface with oxidation. The oxide film at the surface is evaporated simply by heating the substrate in ultra-high vacuum.

To confirm the evaporation and to see the growth mode during the growth, we need some in-situ monitor of the surface state. For that purpose a conventional method is refractive high energy electron diffraction, RHEED. In the RHEED technique, as illustrated in Fig.1.8(a), electron beam with 15~30keV acceleration is injected onto the surface with very shallow angle and the diffraction pattern of reflected beam is imaged on the illumination screen. The image reflects the atomic structure of the surface, that is, it is the pattern of reciprocal

lattice. Because the incident beam goes onto the surface with very shallow angle, when the surface is a mirror, the diffraction is close to two-dimensional, that is, the diffraction pattern is a set of vertical reciprocal lattice “rods”. The image on the screen is a slice of the reciprocal rods with a plane almost parallel to the rods. Actual diffraction patterns of rods have some widths due to various reasons and in such a two-dimensional growth, images like upper-left of Fig.1.9 are obtained.

The image in Fig.1.9 has a strong diffraction spot at upper-center. This is due to the simple mirror-like reflection from the surface (mirror spot) and the more flat is the surface, the higher the intensity is. After opening the shutters molecular beams reach the substrate and the growth starts. Molecules or atoms migrate on the surface of the substrate with thermal activation after adsorption and hit the lattice points at last, forming strong bonding to substrate crystals. This is one of the possible mechanisms for crystal growth and such a “state of growth” is called “growth mode”. The growth mode mentioned above is called layer-by-layer mode.

In the initial stage of layer-by-layer mode, one atomic monolayer growth contains a cycle from a flat surface through a rough surface to a new flat surface. Such a single cycle causes one period in intensity oscillation of mirror spot. The oscillation hence makes it possible for us to monitor the growth of each atomic layer. The oscillation damps in proceeding of the growth due to some incoherency though in many cases growth interruption recovers the flatness for lowering the surface energy of roughness. These properties opens up a way to “flat surface growth” of “migration enhanced epitaxy”, in which the intensity of mirror spot is monitored and the shutters are controlled to keep the highest intensity in the oscillation.

With increasing substrate temperatures, generally the dominant growth mechanism changes into “step flow mode”, in which migrating atoms on the surface attach to the edges of surface steps causing widening of terraces, that is, flow of steps. In this mode no oscillation occurs.

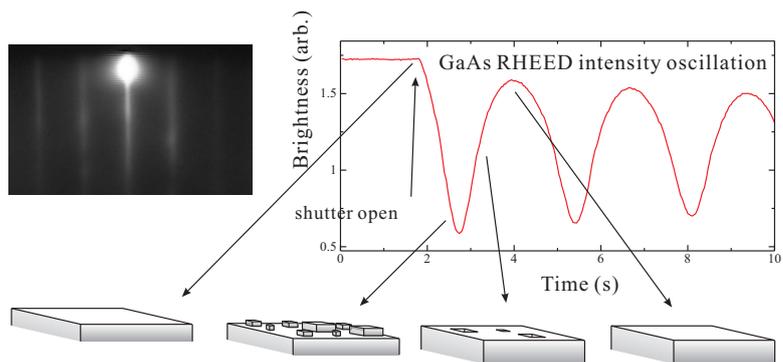


Figure 1.9: Upper-left: RHEED image for two-dimensional MBE growth. The intense spot at upper-center is the mirror spot. The right shows the oscillation of the intensity at the mirror spot as the film grows. The cartoons illustrate the surface states at indicated points in the intensity oscillation of the mirror spot.

## References

- [1] N. W. Ashcroft and N. D. Mermin, “Solid State Physics” Chapter 4 (Brooks/Cole Publishing, 1976).
- [2] D F Johnston, Rep. Prog. Phys., **23**, 66 (1960).
- [3] D. K. Ferry, “Semiconductors: Bonds and bands” (IOP Publishing, 2013).

## Chapter 2 Band structure, effective mass approximation

In solid-state physics, the term “band structure” refers to the energy dispersion relations of the crystal eigenstates in the reciprocal lattice space introduced in the previous chapter. The theme of this chapter is introduction of the concept and how to calculate it theoretically. In addition, we will introduce the effective mass approximation, which is indispensable for handling band electrons in a simple and clear view.

### 1 Band electrons

In free space, the kinetic energy of an electron takes a continuous value from zero. On the other hand, the kinetic energy takes a discrete value in the bound state in the localized potential of the nucleus. There are two views on the energy eigenstates in the periodic potential. One is the perturbation to the state in free space, which creates a section (energy gap) where the eigenvalues do not exist, and the energy eigenvalues are cut to bands. The other is that the discrete level due to the localized potential spreads in a band due to the tunnel between the adjacent sites. The former is called **nearly free electron approximation** (NFEA), and the latter is called **tight-binding approximation** (TBA).

#### 1.1 Bloch theorem

It is needless to prove the Bloch theorem, which is very basics of the solid state physics. For the reference, the theorem is described as follows. Energy eigenstates in a periodic potential are expressed in the real space ( $\mathbf{r}$ ) expression as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (1.1)$$

where  $n$  is the band index,  $u_{n\mathbf{k}}$  is a function with the lattice periodicity, *i.e.*

$$\forall \mathbf{R} \in \{(\text{lattice vector})\}, u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}). \quad (1.2)$$

Here  $\mathbf{k}$  is the wavenumber.

#### 1.2 Nearly free electron approximation (NFEA)

We write the equation for the eigenstates in a lattice potential as

$$\mathcal{H}\psi(\mathbf{r}) = \left[ -\frac{\hbar^2}{2m_0} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (1.3)$$

where  $V(\mathbf{r})$  is the lattice potential.

From the periodicity of  $V(\mathbf{r})$ ,  $u_{\mathbf{k}}(\mathbf{r})$ , they can be Fourier expanded as

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}}, \quad u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}}, \quad (1.4)$$

where  $\mathbf{G}$  are the reciprocal lattice vectors. With substituting (1.1) and (1.4) into the Schrödinger equation (1.3), we obtain

$$\sum_{\mathbf{G}} \left[ \left\{ \frac{\hbar^2}{2m_0} (\mathbf{k} + \mathbf{G})^2 - E \right\} C_{\mathbf{G}} + \sum_{\mathbf{G}'} V_{\mathbf{G}-\mathbf{G}'} C_{\mathbf{G}'} \right] e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} = 0.$$

Because each term in the sum of  $\mathbf{G}$  should be zero, the following simultaneous equations for  $\{C_{\mathbf{G}}\}$  are obtained.

$$\sum_{\mathbf{G}'} \left[ \left\{ \frac{\hbar^2}{2m_0} (\mathbf{k} + \mathbf{G})^2 - E \right\} \delta_{\mathbf{G}\mathbf{G}'} + V_{\mathbf{G}-\mathbf{G}'} \right] C_{\mathbf{G}'} = 0. \quad (1.5)$$

The condition for eq.(1.5) to have non-trivial solutions is

$$\left| \left[ \left\{ \frac{\hbar^2}{2m_0} (\mathbf{k} + \mathbf{G})^2 - E \right\} \delta_{\mathbf{G}\mathbf{G}'} + V_{\mathbf{G}-\mathbf{G}'} \right]_{\mathbf{G}\mathbf{G}'} \right| = 0. \quad (1.6)$$

In NFEA, we consider the perturbation  $\delta V_{\mathbf{G}-\mathbf{G}'}$  to ( $V(\mathbf{r}) = 0$ )

$$\psi(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}}, \quad C_0 = 1, \quad C_{\mathbf{G}} = 0 \quad (\mathbf{G} \neq 0), \quad E = \frac{\hbar^2 \mathbf{k}^2}{2m_0}. \quad (1.7)$$

As a result of perturbation,  $\delta C_{\mathbf{G}}$  is caused. In (1.5), the terms  $\delta V\delta C$ ,  $\delta E\delta C$  are in the higher order to be ignored. Then

$$\frac{\hbar^2}{2m_0}[(\mathbf{k} + \mathbf{G})^2 - \mathbf{k}^2]\delta C_{\mathbf{G}} + V_{\mathbf{G}} = 0 \quad \therefore \delta C_{\mathbf{G}} = \frac{2m_0}{\hbar^2} \frac{-V_{\mathbf{G}}}{(\mathbf{k} + \mathbf{G})^2 - \mathbf{k}^2}.$$

However the approximation collapses at

$$(\mathbf{k} + \mathbf{G})^2 - \mathbf{k}^2 = 0. \quad (1.8)$$

Therefore around the point (1.8), we approximate that only  $C_0$  and  $C_{\mathbf{G}}$  are non-zero. Then we can write down (1.6) as

$$\begin{vmatrix} \frac{\hbar^2}{2m_0}\mathbf{k}^2 - E & V_{-\mathbf{G}} \\ V_{\mathbf{G}} & \frac{\hbar^2}{2m_0}(\mathbf{k} + \mathbf{G})^2 - E \end{vmatrix} = 0, \quad (1.9)$$

which gives the energy eigenstates as

$$E = \frac{1}{2}[E^{(0)}(\mathbf{k}) + E^{(0)}(\mathbf{k} + \mathbf{G})] \pm \frac{1}{2}\sqrt{[E^{(0)}(\mathbf{k}) - E^{(0)}(\mathbf{k} + \mathbf{G})]^2 + 4|V_{\mathbf{G}}|^2}, \quad (1.10)$$

where  $E^{(0)}(\mathbf{k}) \equiv \hbar^2\mathbf{k}^2/2m_0$ . The result indicates the appearance of the energy separation of  $\pm V_{\mathbf{G}}$  (**bandgap** or **forbidden band**). For a system with the lattice constant  $a$ , the condition (1.8) is  $2a \cos \theta = n\lambda$  ( $n$  is an integer,  $\lambda$  is the wavelength of electron). This is nothing but the Bragg condition for diffraction of waves. Thus the result can be interpreted as the electron wave get a Bragg reflection from the lattice and the interference between the waves creates a standing wave, which results in the bandgap.

### 1.3 Reduced zone expression

A Bloch function can be written as follows with  $\mathbf{G}$  a reciprocal lattice vector as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} = u_{n\mathbf{k}}(\mathbf{r})e^{-i\mathbf{G}\cdot\mathbf{r}}e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}.$$

Function  $v(\mathbf{r}) \equiv u_{n\mathbf{k}}(\mathbf{r})e^{-i\mathbf{G}\cdot\mathbf{r}}$  also has the periodicity  $v(\mathbf{r}) = v(\mathbf{r} + \mathbf{R})$ , where  $\mathbf{R}$  is lattice vectors.  $\psi_{n\mathbf{k}}$  can thus be expressed as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \xi_{n'\mathbf{k}+\mathbf{G}}(\mathbf{r}) \quad (1.11)$$

with another Bloch function  $\xi_{n'\mathbf{k}}$ . Namely, the expression in (1.1) has the arbitrariness of reciprocal lattice vectors. In other words, when a wavefunction has a spatial modulation of lattice period, there is ambiguity whether the modulation is included in the lattice periodic part  $u(\mathbf{r})$  or in the plane wave part  $e^{i\mathbf{k}\cdot\mathbf{r}}$ .

On the other hand, the system represented by Schrödinger equation (1.3) has the time-reversal symmetry and  $E(\mathbf{k}) = E(-\mathbf{k})$ . The above two relations on  $E(\mathbf{k})$  leads to  $E(\mathbf{G} + \mathbf{k}) = E(\mathbf{G} - \mathbf{k})$ . That is,  $E(\mathbf{k})$  is symmetric to the zone boundaries.

The arbitrariness in (1.11) leads to the arbitrariness in the representation of  $E(\mathbf{k})$ . As Fig.1.2(a), in **extended zone representation**,  $E(\mathbf{k})$  is represented as a single-valued function of  $\mathbf{k}$  while as in Fig.1.2(b), in **reduced zone representation**, representation of  $E(\mathbf{k})$  is folded into the first Brillouin zone.

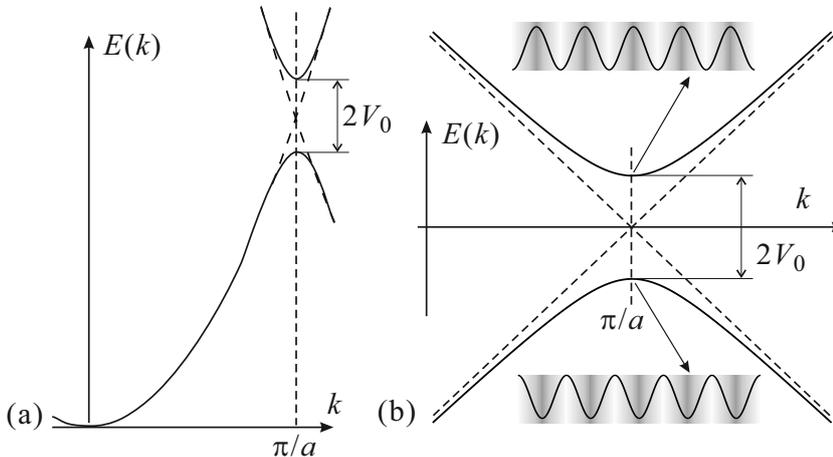


Figure 1.1: (a) In NFEA, a bandgap of (1.10) appears at  $k = G/2$ . (b) Blowup of the region around the bandgap in figure (a). The cartoons explain why the standing waves get the energy gap

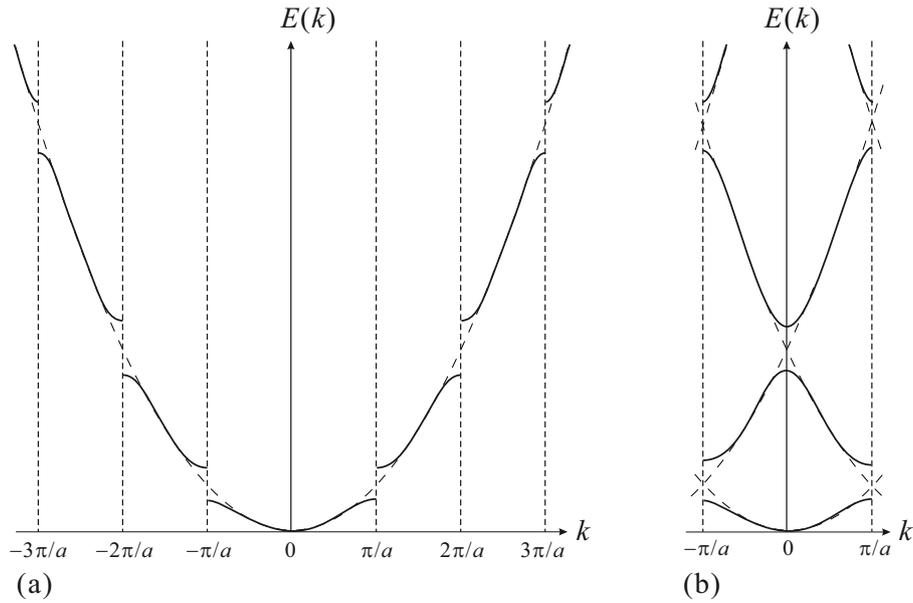


Figure 1.2: Two ways of expression for the energy bands of NFEA. (a) Extended zone expression. (b) Reduced zone expression.

distance :	reciprocal lattice	number of points
0 :	(0,0,0)	1
$\sqrt{3}$ :	(1,1,1), (1,1,-1), (1,-1,1), ...	8
2 :	(2,0,0), (0,2,0), (0,0,2), (-2,0,0), ...	6
$\sqrt{8}$ :	(2,2,0), (2,0,2), (0,2,2), (-2,2,0), ...	12
$\sqrt{11}$ :	(3,1,1), (1,3,1), (1,1,3), (-3,1,1), ...	24
.....		

Table 2: Classification of reciprocal lattice points with the distance from  $\Gamma$ -point (unit  $G_0 \equiv 2\pi/a$ ).

In Fig.1.2(b), the second and the third bands is obtained from cutting for the first Brillouin zone  $[-\pi/2, \pi/2]$  of the overlap of the extended zone representations from the two neighboring reciprocal lattice points  $k = \pm 2\pi/a$ . This is a natural consequence that the reduced zone representation is possible from the arbitrariness of the reciprocal lattice vectors as in (1.11).

## 1.4 Empty lattice approximation

In NFEA, in the limit of  $V_0 \rightarrow 0$ , this is nothing but a free space and the energy gap disappears, the dispersion relation is simply parabolic. However, the free space is not a system in which the spatial periodicity of the lattice is lost, and the continuous translational symmetry also includes the periodic translational symmetry of the lattice. Rather, it can be considered that the lattice of the empty primitive cell remains. In the **empty lattice approximation**, we hence consider the reduced zone representation of the parabolic energy dispersion. In Bloch function representation, the plane wave function  $e^{i\mathbf{k}' \cdot \mathbf{r}}$  is separated into the lattice periodic part  $u_{n\mathbf{k}}(\mathbf{r})$  and the crystal wavenumber part  $e^{i\mathbf{k} \cdot \mathbf{r}}$  and apply reduced zone representation.

Let's take an example with a three-dimensional crystal. Consider the reciprocal lattice and Brillouin zone in Fig.1.4 in the case of fcc. First, to obtain the reduced zone representation, since the principle of reduced zone representation is indefiniteness of the reciprocal lattice vector as in (1.11), consider the reciprocal lattice of fcc and the bcc lattice of Fig.1.4 (a), and draw parabola with the origin at each reciprocal lattice point. Then we cut the diagrams with the first Brillouin zone shown in Fig.1.4(c). The reciprocal lattice points we need to consider in this drawing are summarized in Tab.2. The farther from the origin, the higher the energy branch of the parabola from the reciprocal lattice point.

The problem with a three-dimensional band structure expression is how to display it. It is not possible to draw multiple parabolas in a three-dimensional space. Usually we only draw the energy dispersion on some representative lines in the reciprocal space. Fig.1.3 shows a way of drawing often used to display the band structure. The energy dispersions on the lines which connect points with high symmetries are drawn. As in

the figure the line goes along  $L \rightarrow \Gamma \rightarrow X \rightarrow K \rightarrow \Gamma$ . (a) shows the empty lattice approximation while (b) shows the realistic band dispersion in Si calculated with empirical pseudo-potential approximation (we will see in the next week). There is no bandgap in the empty lattice approximation naturally. On the other hand, we see clear resemblance between them. When we go into realistic calculations with gaps, the diagram is useful to see which branch corresponds to which reciprocal point. Furthermore, when a level repulsion causes energygap, we need to consider symmetry of the lattice and the empty lattice approximation is also useful for seeing that.

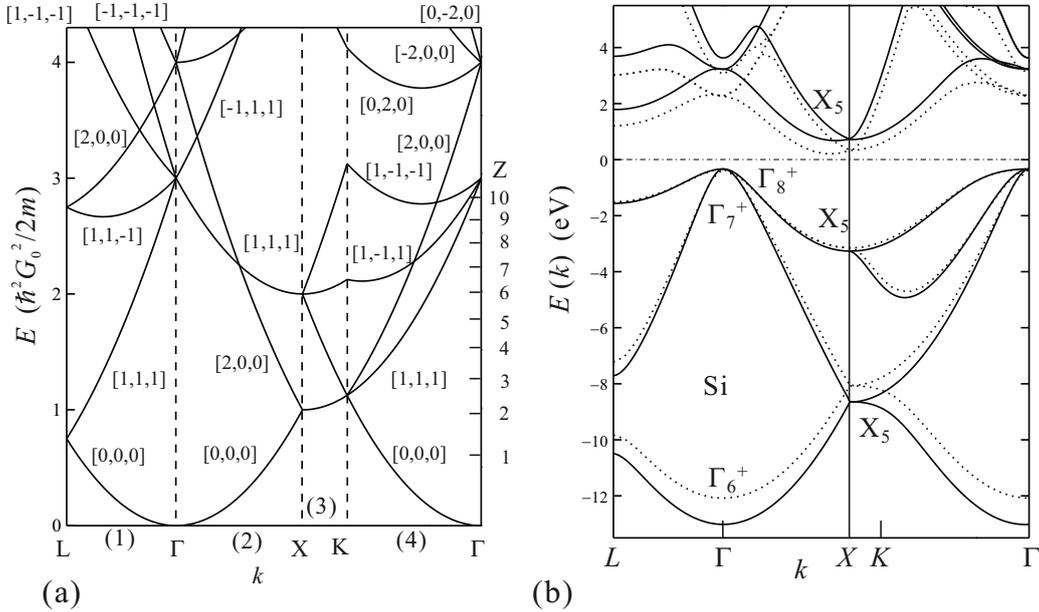


Figure 1.3: (a) Empty lattice approximation of fcc-lattice. Three numbers in  $[\dots]$  indicate corresponding origins of parabolas. (b) Realistic band structure of Si calculated by empirical pseudo-potential method.

## 1.5 Tight binding approximation

In the next week we begin with tight-binding approximation.