Lecture note on Physics of Semiconductors (8)

2nd June (2021)   Shingo Katsumoto, Institute for Solid State Physics, University of Tokyo

Next we see FETs without pn-junction. For transistor action, they utilize phenomena on the surfaces or interfaces. In homo-type pn-junctions the uniformity of space is broken by impurity doping. They do not use interfaces or surfaces. This was important for Shockley and co-workers to realize "stable and reproducible" devices because for the semiconductor technologies in those days control of surfaces or interfaces was too difficult for commercial production. Even the high quality crystalline growth and the accurate doping technique, which are indispensable for the realization of pn-junctions, were surprisingly high technique. However the great strides in semiconductor technologies caught the control techniques of surfaces and interfaces in incredibly short time. Naturally there were movement to utilize them for device actions and they overwhelmed bulk shortly. We have a look for these representative modern devices here. But the limit of miniaturization inevitably requires three dimensionality nowadays and we do not know what happens next.

6.3.3   Schottky barrier (junction)

Here we consider junctions between semiconductors and metals. Simple guiding principles are

1. Rigid band approximation,
2. Recovery of bulk states away from the junction,
3. In equilibrium E_{F} (μ) is constant over the space.

On semiconductor surfaces, there usually are **surface states** with high density of states. Metal-semiconductor junctions are strongly affected by those states. Here, however, we first look what Anderson's rule tells about the interface[**?**]. The baseline of rigid bands can be taken to an edge of "band", in which electrons can freely travel between the metal and the semiconductor. It is usually impossible to find such an energy band inside insulators and semiconductors, which have very different energy bands. Then such a "band" can be found as the vacuum levels. Then the excitation energy required is so called **work function**. Let the work functions in the semiconductor and the metal $e\phi_{\mathrm{S}}$ and $e\phi_{\mathrm{M}}$ respectively. Generally $e\phi_{\mathrm{M}} \neq e\phi_{\mathrm{S}}$. On the other hand, from the guiding principle 2., the bulk E_{F}'s in the metal and in the semiconductor away from the junction should be the same. And E_{F} should be constant throughout the system.

The following procedure, of course, is not real physical process but just a virtual process inside human brain, for construction of consistent band alignment. The final result, however, may be realized in the model of junctions though there still remain many idealizations and reality should be much more complex.

We assume $e\phi_{\mathrm{M}}$ is larger than $e\phi_{\mathrm{S}}$, the semiconductor is doped to n-type and the donor concentration is N_{D}. We make the vacuum levels in the both sides fit to each other and extrapolate the bulk band structures to the interface to obtain the band alignment shown in Fig. 6.9(a). Here the Fermi level in the semiconductor places higher than that in the metal causing flow of carriers from the semiconductor to the metal. The carrier flow generates charge accumulation at the interface creating an electric field perpendicular to the junction plane. The metallic side is also charged up but it has much higher charge concentration, which screens the electric field within the screening length less than a lattice constant making the band bending negligible in this side. Let the accumulated charge in the metal side per unit area $-Q$, in the semiconductor side ($x > 0$, interface at $x = 0$), the electric field at x is $(eN_D x - Q)/\epsilon\epsilon_0$ and the potential difference between 0 and x_d is

$$\phi(x_d) = \int_0^{x_d} (eN_D x - Q)/\epsilon\epsilon_0 \, dx = \frac{1}{\epsilon\epsilon_0} \left(\frac{eN_D}{2} x_d^2 - Q x_d \right). \tag{6.30}$$
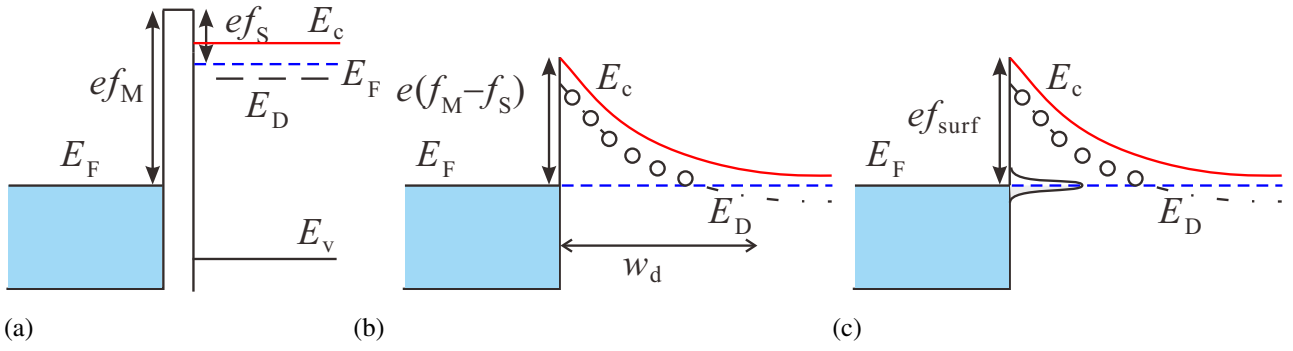
Fig. 6.9 (a) Virtual band alignment, in which a metal and a semiconductor are connected as the vacuum levels for them agree. (b) Band bending effect to make E_F constant throughout the junction is superposed to the alignment in (a). The situation corresponds to an ideal interface without surface states at the semiconductor side. (c) Illustration of Fermi level pinning by surface states. The surface potential ϕ_{surf} is determined by the position of the dominant surface states from the band edge E_c. This usually has nothing to do with the difference between the work function.

Let the space charge (depletion) layer width be w_d. The condition that electric field outside the depletion layer should be zero, gives $w_d = Q/eN_D$. On the other hand, the condition $e\phi(w_d) = \phi_M - \phi_S$ also gives Q as

$$Q = \sqrt{2\epsilon\epsilon_0 N_D e(\phi_M - \phi_S)}, \quad \therefore w_d = \sqrt{\frac{2\epsilon\epsilon_0(\phi_M - \phi_S)}{eN_D}} \equiv \sqrt{\frac{2\epsilon\epsilon_0 V_s}{eN_D}}. \tag{6.31}$$

Here we write $eV_s \equiv \phi_M - \phi_S$. Now we can illustrate the band structure for electrons (holes for p-type) around the metal-semiconductor interface as in Fig. 6.9(b), showing a potential barrier, which is called **Schottky barrier**.

An external voltage V is mostly bared in the semiconductor side, and the height of the barrier changes to $e(V_s - V)$ while the height from the metal side remains as eV_s. To be more accurate, we need to consider the kinetic energy distribution in the semiconductor and count the number of electrons which go over the barrier. But here for simplicity we assume the kinetic energy of electrons in the semiconductor is a constant. Then the equation for thermal electron emission from metallic surface can be applied to obtain

$$J = AT^2 \left[\exp\left(\frac{e(V - V_s)}{k_B T}\right) - \exp\left(\frac{-eV_s}{k_B T}\right) \right] = eAT^2 \exp\left(\frac{-eV_s}{k_B T}\right) \left[\exp\left(\frac{eV}{k_B T}\right) - 1 \right]. \tag{6.32}$$

Here A is the **Richardson coefficient**. The first term is current from the semiconductor side, the second is that from the metal side. The current-voltage characteristics is similar to that of a pn-junction with the Schttkey barrier height corresponding to the built-in potential.

In the above the surface of semiconductor is too much idealized for it to have no surface states. However in real metal-semiconductor junctions, current-voltage characteristics are similar to eq.(6.32). One big difference is in eq.(6.32), the barrier height should change with changing the metal species but in reality, the barrier height is almost constant for semiconductor species and independent of metals. This is due to the **surface states** on the semiconductors. The surface states have narrow energy widths, very high density of states pinning the Fermi level to the center of them. Hence the band bending exists even before the connection to metals and the alignment is accomplished between the metal E_F and the surface states. This is called **pinning of Fermi level** by the surface states.

Once the Fermi level is pinned by the surface states, the band bending is determined by semiconductor species. Hence when n-type Schottky barrier can be formed for a semiconductor for example, p-type is not available for the same semiconductor. The other way around. Actually, for GaAs, p-type Schottky barrier is not available while for InP, n-type Schottky barrier is difficult. This makes it difficult to obtain complementary devices which utilize Schottky barriers. In the case of metal-oxide-semiconductor (MOS) devices, an **inversion layer** formed by *e.g.*, pushing down a band of a p-type semiconductor and turning it to an n-type channel, can be used for complementary device. This is, however, impossible for Schottky devices.

6.3.4 MES-FET

Among III-V semiconductors, GaAs is frequently used for electric devices as well as for optical devices. But it is difficult to form good quality oxide layers on the surfaces, hence no MOS type device for GaAs is available. Instead, MEtal-Semiconductor FET (MES-FET) structure has been frequently adopted. GaAs has light electron mass, high mobilities. And the effective capacitance of Schottky diode can be small. Hence GaAs MESFETs are often used for high-frequency application.
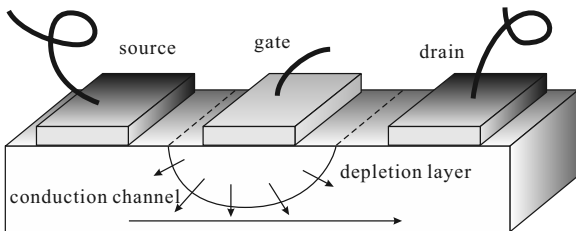
As shown in the left figure, the structure of MES-FET is simple. The conduction channel thickness is controlled with the reverse bias voltage (**gate voltage**) through that of depletion layer. The device action, characteristics are similar to those for JFET. Schottky junctions have larger leak current in gate characteristics, only single carrier type is available and complementary circuits cannot be composed with them. These properties are great obstacles for large scale integration. MES-FETs are still widely used as high frequency devices for *e.g.*, microwave.

6.3.5 MOS structure

As named, a thin oxide film for insulation is inserted between a metal and a semiconductor in a Metal-Oxide-Semiconductor (MOS) structure. Needless to say, most frequently used Si has SiO_2 as the oxide layer, which is very stable and has good insulation characteristics. An SiO_2 film can be easily formed with thermal oxidation onto a Si. Both p-type and n-type channels can be controlled and Complementary MOS (CMOS) circuits are easily realized. Also with low gate leakage current, high on-conductance, off-resistance, the power consumption in logic circuits jumped down with the CMOS circuits hence increased degree of integration. Now CMOS is doubtlessly the king of semiconductor circuits. A few decades ago high speed logic circuits were mainly composed with Emitter Coulpled Logic (ECL) of BJT but the requirement of large scale integration and the increase of cut-off frequency in CMOS circuit have made drastic change and now, even so called supercomputers are using CMOS circuit in CPU.

MOSFET structure also resembles to JFET and the essential difference to MESFET is the existence of thin oxide layer between the semiconductor and the gate metal. In a **depletion type** device, the conduction channel is pinched by depletion layer while in a enhancement type device, the band is pushed down with gate electric field to form conduction channel. An oxide layer bears much higher voltage than a Schottky barrier, hence with a strong bending, *e.g.*, formation of an n-type two-dimensional conduction channel below a p-type semiconductor surface (**inversion layer**).

Si-MOS structures are now used not only in integrated circuits but also for power devices. Recently however, SiC is collecting wider interest for power devices because of the lower ON-resistances. And for high-frequency power devices,
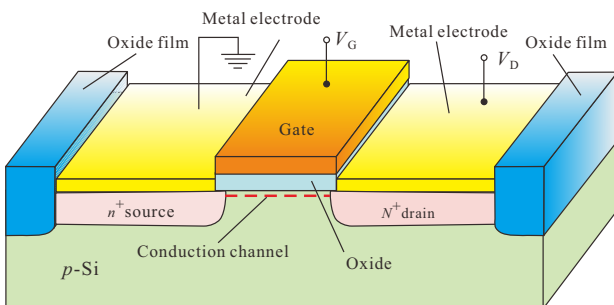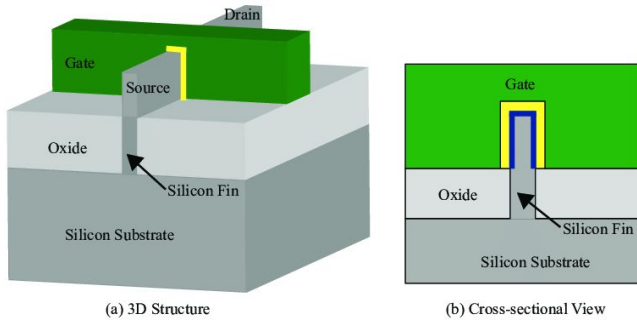

Fig. 6.10 Schematic view of a MOSFET device. In fabrication holes are opened on thermally oxidized films with lithography. The dopants are diffused through the holes. The structure like this often appears due to the process.

the weight of developement is shifting to GaN-related materials.

(a) 3D Structure

(b) Cross-sectional View

6.3.6 FinFET

In the beginning of 21st century, in semiconductor integrated circuits, fierce competition for improving the degree of integration continued, centered on CMOS. Compared to other logic circuit schemes, CMOS is overwhelmingly advantageous in terms of power consumption. Even for such CMOS scheme, what limits the degree of integration is in-chip heating due to power consumption. In order to solve this, device driving at low voltage has come to be required. A way is to replace SiO_2 with some other insulating thin film that has higher dielectric constant. With this, ON/OFF action of the channels would be available. The letter κ is often used as the symbol of the dielectric constant and such dielectrics are called "high-κ" or simply "high-k" materials. From various restrictions, now hafnium silicate, hafnium oxide, zirconia are used for such high-κ materials.

Furthermore, the FinFET, in which the channel shape is changed from the planer type to the fish-fin shaped, has been now widely used. As illustrated in the figure, in a FinFET, a thin channel is covered with the gate and the depletion layer/ the inversion layer grow over the channel from both sides of the "fin", resulting in faster switching rate (less than 1 ps) and higher ON-conductance than those of planer structure. Also the device density can be higher. Now they are the main structure for the logic LSIs.

6.4 Heterojunction

The two materials on the sides of a junction have similar properties, lattice structures, etc. to each other in semiconductor **heterojunction** in comparison with Schottky junctions or with MOS structures. As a result, in semiconductor heterojunctions, sharp changes in the effective potential can be realized, the quantum coherence of electrons is kept over the junctions. Therefore they can be used for the building block of the devices which utilize quantum effects such as electron tunneling. And with heterojunctions, one even can create new periodic structures in solids and modify the band structure. This is called **band engineering**.

Because this lecture is for the physics in semiconductors, we begin with how to treat such hetero-interface physically.

6.4.1 Effective mass approximation at hetero-interfaces

As in the textbook[1] written by myself or that by Bastard[2], let us consider hte application of effective mass approximation for a hetero-interface under simplest case.

6.4.1.1 Hetero-interface

We consider the situation in which semiconductors A and B (A : $z < 0$, B : $z > 0$) are connected at $z = 0$ (xy-plane). In each region, the Bloch theorem is appled to write

$$\psi^{(A)}(\boldsymbol{r}) = \sum_l f_l^{(A)}(\boldsymbol{r}) u_{l\boldsymbol{k}}^{(A)}(\boldsymbol{r}), \quad \psi^{(B)}(\boldsymbol{r}) = \sum_l f_l^{(B)}(\boldsymbol{r}) u_{l\boldsymbol{k}}^{(B)}(\boldsymbol{r}), \tag{6.33}$$

where l is the band index, $u_{l\boldsymbol{k}}^{(A,B)}$ are functions with the lattice periodicity. For simplicity, the lattice periodic part of the Bloch functions and the band dispersions are the same other than the positions of band bottoms and tops.

$$u_l^{(A)}(\boldsymbol{r}) = u_{l\boldsymbol{k}}^{(B)}(\boldsymbol{r}), \quad \partial\epsilon_l^{(A)}/\partial\boldsymbol{k} = \partial\epsilon_l^{(B)}/\partial\boldsymbol{k}.$$

With this simplification, the continuity condition of wavefunction at $z = 0$ gives

$$f_l^{(A)}(\boldsymbol{r}_{xy}, 0) = f_l^{(B)}(\boldsymbol{r}_{xy}, 0),$$

where $\boldsymbol{r}_{xy}$ is a vector in the xy-plane. For the freedom of $\boldsymbol{r}_{xy}$, the Bloch theorem tells

$$f_l^{(A,B)} = \frac{1}{\sqrt{S}}\exp(i\boldsymbol{k}_{xy}\cdot\boldsymbol{x})\chi_l^{(A,B)}(z),$$

where $1/\sqrt{S}$ is the partial normalization factor of plane wave in xy-plane, $\chi_l(z)$ is the envelopefunction along z-direction.

For the freedom along z-direction, we consier the $k\cdot p$ perturbation. That is, first we obtain the lattice periodic function and the discrete levels for $k = 0$ and the wavefunctions for $k \neq 0$ are obtained by the hybridization of these wavefunctions caused by the perturbation Hamiltonian, which is proportional to $k\cdot p$. We write down the equation for $\boldsymbol{\chi} = \{\chi_j\}$ as

$$\mathscr{D}^{(0)}\left(z, -i\hbar\frac{\partial}{\partial z}\right)\boldsymbol{\chi} = \epsilon\boldsymbol{\chi}, \tag{6.34}$$

where the $N \times N$ matrix of operators $\mathscr{D}^{(0)}$ is

$$\mathscr{D}_{lm}^{(0)}\left(z, \frac{\partial}{\partial z}\right) = \left[\epsilon_l(z) + \frac{\hbar^2 k_{xy}^2}{2m_0} - \frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial z^2}\right]\delta_{lm} + \frac{\hbar\boldsymbol{k}_{xy}}{m_0}\cdot\langle l|\boldsymbol{p}_{xy}|m\rangle - \frac{i\hbar}{m_0}\langle l|p_z|m\rangle\frac{\partial}{\partial z} \tag{6.35}$$

with

$$\epsilon_l(z) = \epsilon_l^{(A)} \quad (z < 0), \quad \epsilon_l^{(B)} \quad (z \geq 0). \tag{6.36}$$

Here we write $|u_{m0}\rangle$ as $|m\rangle$, etc.

Emphasizing "band-discontinuity potential," we write

$$V_l(z) \equiv \begin{cases} 0 & z < 0 \quad (z \in A) \\ \epsilon_l^{(B)} - \epsilon_l^{(A)} & z \geq 0 \quad (z \in B). \end{cases} \tag{6.37}$$

Then we reach the simultaneous equation of $\{\chi_l\}$ as [1]

$$\sum_{m=1}^{N}\left\{\left[\epsilon_{m0}^{(A)} + V_m(z) + \frac{\hbar^2 k_{xy}^2}{2m_0} - \frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial z^2}\right]\delta_{lm} - \frac{i\hbar}{m_0}\langle l|\hat{p}_z|m\rangle\frac{\partial}{\partial z} + \frac{\hbar\boldsymbol{k}_{xy}}{m_0}\cdot\langle l|\hat{\boldsymbol{p}}_{xy}|m\rangle\right\}\chi_m = \epsilon\chi_l. \tag{6.38}$$

Let us consider the continuity condition of the envelope function χ_l of band l. Because we have assumed that u_l is common for A and B, χ_l should be continuous at the interface. On the other hand, the integration of (6.38) over the interface and the continuity of χ_l leads to the condition

$$\mathscr{A}^{(A)}\chi^{(A)}(z_0 = 0) = \mathscr{A}^{(B)}\chi^{(B)}(0), \tag{6.39}$$

where

$$\mathscr{A}_{lm} = -\frac{\hbar^2}{2m_0}\left[\delta_{lm}\frac{\partial}{\partial z} + \frac{2i}{\hbar}\langle l|p_z|m\rangle\right]. \tag{6.40}$$

It is now clear that the band-hybridizing term $\langle l|p_z|m\rangle$ from the $k\cdot p$ perturbation breaks the simple continuity of derivative of the envelope function.

---

[1] If we go up to the second order in k, we have many other terms, which makes the equation very complicated. We thus have omitted them.

6.4.1.2 Joint of envelope function

Next we do not equate u nor the band dispertion (effective mass) but only the single band is considered. The effective mass equation is a second-order differential equation, and the general boundary connection conditions are as follows.

$$\begin{pmatrix} \chi^{(A)}(0) \\ \nabla_A \chi^{(A)}(0) \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} \chi^{(B)}(0) \\ \nabla_B \chi^{(B)}(0) \end{pmatrix}, \tag{6.41}$$

where, taking a as the common lattice constant,

$$\nabla_{A,B} = \frac{m_0}{m_{A,B}} \frac{\partial}{a\partial z}. \tag{6.42}$$

$T_{BA} = \{t_{ij}\}$ is called **interface matrix**.

The particle current density along z is determined by the envelope function as

$$j(z) = \frac{\hbar}{2im^*} \left[\chi^*(z) \frac{\partial \chi}{\partial z} - \frac{\partial \chi}{\partial z} \chi(z) \right]. \tag{6.43}$$

From the particle-number conservation, $j(z)$ in A and B regions should be the same. The condition is equivalent to

$$\det T_{BA} = 1. \tag{6.44}$$

Because this condition is fulfilled when T_{BA} is the unit matrix I, the simplest envelope function approximation is to put $T_{BA} = I$. In this case, the envelope function can be treated just the same as the real wavefunction. In the case of GaAs-(Al,Ga)As interface, the interface matrix obtained for a one-dimensional tight-binding model indicates the envelope function approximation works well.

In such a case, we can consider the step function potential at the boundary with the height of **band discontinuity** which is determined by the combination of the materials. And the envelope function can be viewed as ordinary quantum wavefunction. On the above basis, we now can use methods to design quantum systems such as one-dimensional potential by thin film growth technique.
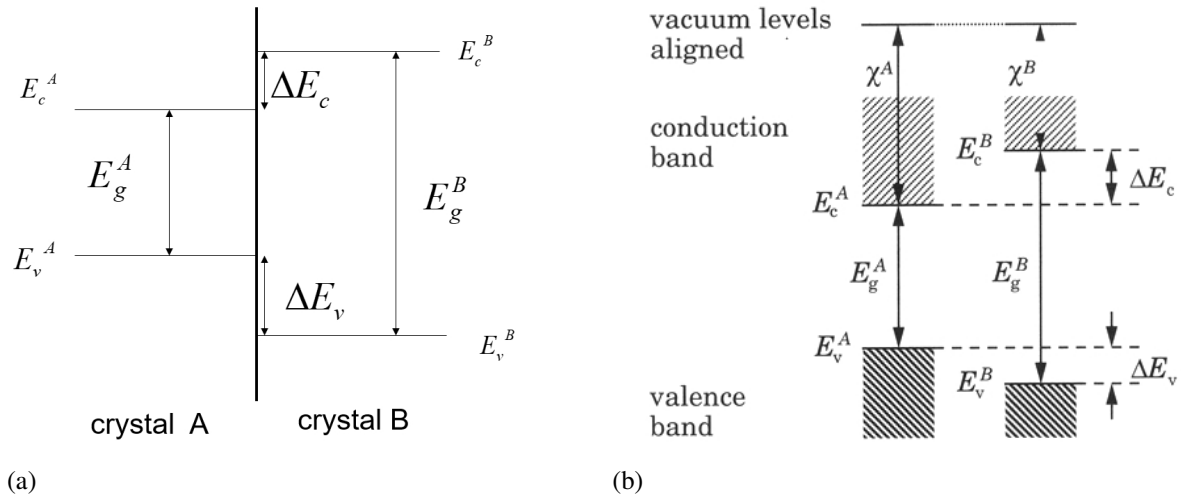

(a)   (b)

Fig. 6.11 (a) Diagram displaying symbols for the band alignment parameters at a junction of crystals A and B. (b) Anderson model, in which the relative positions of bands are determined by the affinities from the vacuum level.

6.4.2 Anderson's model

Figure 6.11 illustrates a long-used Anderson model [*2] for heterojunctions[3]. In the model, as shown in Fig. 6.11(a), the bands in the bulk continue to the interface. The effect of charge transfer is taken into account as built-in potential just like the treatment of pn homo-junctions.

An important point in this model is the relative band position at the heterojunctions. In the Anderson model, as shown in Fig. 6.11(b), this is determined from the quantity called "affinity", which is the lowering in the energy of electrons with condensation into the crystal state. Then in this model, the connection of the bands is determined by the species of the crystals In the figure, the affinities of A and B are χ^A and χ^B respectively.

The model, in itself, has many problems, many of which are on the "affinity." Can the affinities be well-defined? Can we calculate them? Are they measurable? We do not have time to go into the problems and furthermore, the experiments have shown that such simple modeling does not work at the level of device designing, in which we need detailed information of band-discontinuity.

We will have a brief look at the junction types and summarize theoretical approachs to the band-discontinuities in Appendix 6C.

6.4.3 Classification of heterojunctions

Semiconductor heterojunctions are classified phenomenologically by the alignment of bands at the interface. Figure 6.12 shows three types of band alignment. (a) is most frequently found and called type-I. On the larger gap side, the conduction bottom is higher and the valence top is lower. In type-II, as shown in (b), the conduction bottom and the valence top shift to the same direction when an electron passes the interface. There is a common energy gap region for A and B in the case of Fig. 6.12 (b). When the missalignment is larger and the energy gap at the interface is closed as in (c), in Japan they call the alignment type-III and in other countries staggered type=II. For example, in ref. [4], the authors call (c) as type-II.
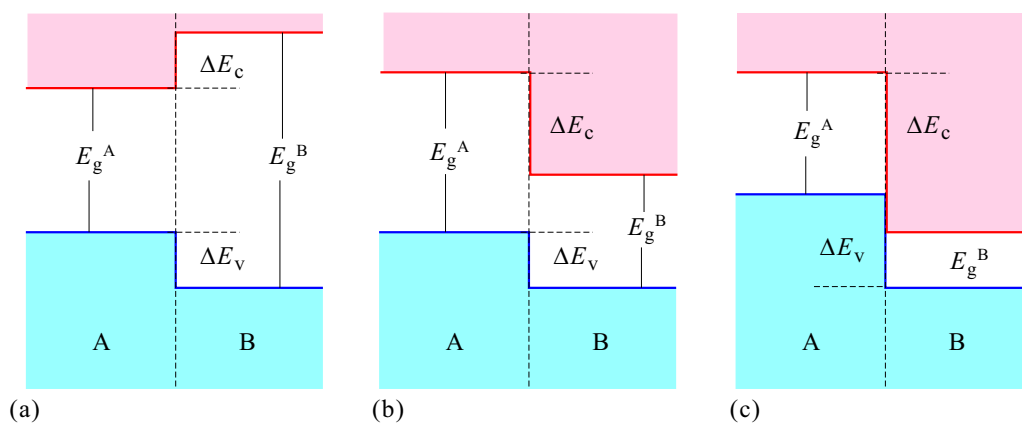

Fig. 6.12 (a) Type-I: To the wider gap semiconductors, the conduction bottom goes up, the valence top goes down. (b) Type-II: The conduction bottom and the valence top moves to the same direction when going through the junction. Also called broken-up or misaligned. (c) Type-III: (Special classification in Japan) Same as type-II but there is no overlapping in the energy gap. Instead there is an overlapp between the valence band on one side and the conduction band on the other side. Also called staggered.

[*2] This "Anderson" is a different person from novel laureate Philip W. Anderson in Bell lab. P. W. Anderson is famous for his "Anderson model" of impurities and R. L. Anderson was in IBM Watson. Bit confusing.

6.5   Formation of heterojunctions

We have already had a look on the epitaxial growth technique. Here I just mention about the lattice matching and energy variation.

6.5.1   Epitaxial growth

Most popular method to form heterojunctions of semiconductors is epitaxial growth already presented in the lecture by Prof. Akiyama. Epitaxial growth methods can be classified into liquid-phase epitaxy, vapour-phase epitaxy, and vacuum deposition. In liquid-phase epitaxy, precipitation onto crystal substrates from melts of ingredients is used. The growths occur in states close to equilibrium and high quality crystals can be obtained while it is hard to obtain sharp interfaces. When one needs sharp interfaces and precise control of layer thicknesses, usually the latter two methods of epitaxy are adopted.

An important point in the formation of heterojunction is the **lattice matching** in lattice constants and crystal systems. In Fig. 6.13, we plot representative compound semiconductors and elemental semiconductors on the plane of lattice constant and energy gap. Most of the plotted semiconductors have a common crystal system, FCC bravais lattice. Vertical gray bands indicate possible groups of lattice matched heterostructure growth though these combinations are not always available in practical growths. Besides these semiconductors, heterojunctions of GaN family are important for industrial demands. They usually have Wurtzite structure (hexagonal close-packed, HCP) and need high temperature treatments, the heterostructures thus are mostly composed within nitride families.

Even with considerable lattice mismatch, a misfit-dislocation free growth to a certain film thickness is possible. An estimation of the thickness given as a balance point of the strain energy concentrated on dislocations and that within whole grown film, is called **Matthews' critical thickness**[5]. Because actual crystal growths are carried out under some non-equilibrium condition, the total free energy not necessarily takes the minimum, the process is generally non-adiabatic. Hence the Matthews' thickness is just a rough estimation. In many cases we need to keep substrate temperatures high
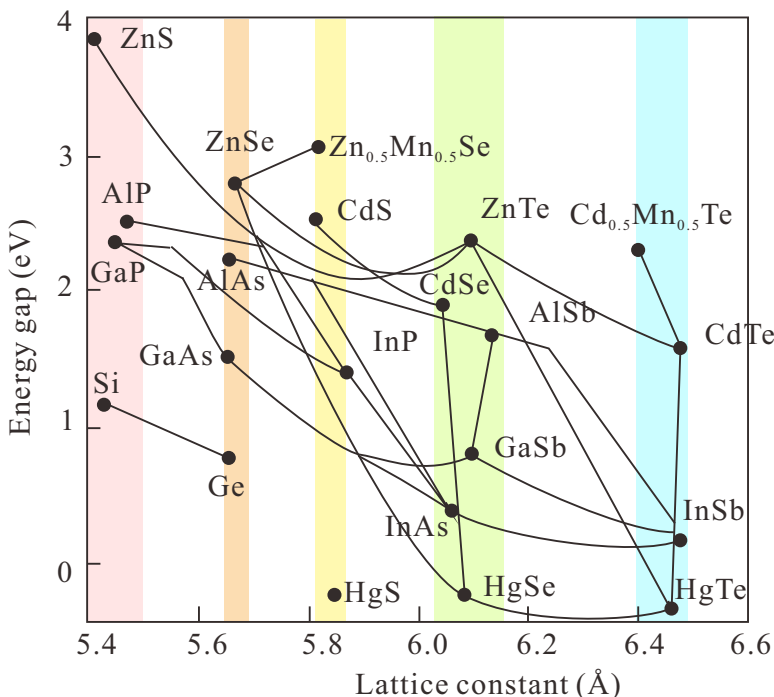

Fig. 6.13   Plots of the lattice constants and the energy gaps of II-VI, III-V compound semiconductors and IV elemental semiconductors. The lines connecting the points indicate possible mixed crystals. Vertical gray bands indicate possible groups of lattice matched heterostructure growth.
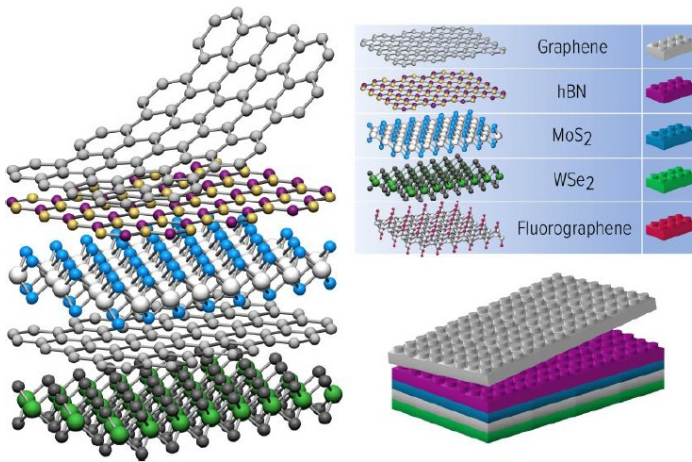
Fig. 6.14 Conceptual illustration of van der Waals heterostructure, which is produced by stacking various two-dimensional materials.

enough during growths and the difference in coefficients of thermal expansion in the two materials sometimes causes dislocations or strains. Many points should be taken into account in actual growths[6].

6.5.2 van der Waals heterostructure

Recently van der Waals heterostructure, which is formed in completely different way, is collecting attentions[7]. That is a mechanical stacking of two-dimensional materials like graphene as shown in Fig. 6.14 (graphene will be introduced later as a two-dimensional electron system without heterointerface). Sometimes epitaxial growth like CVD is adopted but in many cases mechanical stacking of exfoliated two dimensional materials creates high-quality heterostructure, which implies possible completely new formation method of heterostructure.

Chapter 7 Quantum structures (quantum wells, wires, dots)

So far, we have looked at low-dimensional systems such as graphene, which is a two-dimensional substance. These were, so to speak, natural low-dimensional systems. By using heterojunction, metal joining, and microfabrication technology, human hands have come to enter the design of material systems to a certain extent. In this chapter, we look at three typical quantum structures, two-dimensional, one-dimensional, and zero-dimensional systems.

7.1   Quantum well

A region with lower potential sandwiched with two heterojunctions to higher potential materials is **quantum well**. The readers should be familiar with it since introduction of elementary quantum mechanics. In other words, however, the semiconductor heterojunction technology has made the quantum well as a real substance from just an exercise for students.

7.1.1   Discrete quantum levels in a quantum well

Let the well width be L, the barrier height V_0. In $x \leq -L/2$, $L/2 \leq x$ (outside the well) Schrödinger equation is

$$\left[-\frac{\hbar^2 \mathrm{d}^2}{2m \mathrm{d}x^2} + V_0 \right] \psi = E\psi. \tag{7.1}$$

Let us put $\kappa \equiv \sqrt{2m|E - V_0|}/\hbar$ and let $C_{1,2}$, $D_{1,2}$ be constants specific to the regions, the solution outside the well can be written as

$$\psi(x) = \begin{cases} C_1 \exp(i\kappa x) + C_2 \exp(-i\kappa x) & E \geq V_0, \\ D_1 \exp(\kappa x) + D_2 \exp(-\kappa x) & E < V_0. \end{cases} \tag{7.2}$$

In the case of $E < V_0$, the wavefunction should be localized around the well and zero for $x \to \pm\infty$, then

$$L/2 < x \text{ で } D_1^+ = 0, \quad x < -L/2 \text{ で } D_2^- = 0.$$

Superscript $\pm$ distinguish the regions positive/negative of x. Inside the well, letting C_1, C_2 be constants, we write the wavefunction with plane waves as

$$\psi = C_1 \exp(ikx) + C_2 \exp(-ikx), \quad k \equiv \frac{\sqrt{2mE}}{\hbar}, \tag{7.3}$$

where for simplicity, we assume the effective mass m is common for inside and outside the well. The boundary condition at $x = \pm L/2$ where the potential is discontinuous is now applied. Continuity and differentiability at the potential boundary $x = 0$ require

$$\text{Continuity} \begin{cases} C_1 \exp(ikL/2) + C_2 \exp(-ikL/2) = D_2^+ \exp(-\kappa L/2), \\ C_1 \exp(-ikL/2) + C_2 \exp(ikL/2) = D_1^- \exp(-\kappa L/2), \end{cases}$$

$$\text{Differentiability} \begin{cases} ikC_1 \exp(ikL/2) - ikC_2 \exp(-ikL/2) = -\kappa D_2^+ \exp(-\kappa L/2), \\ ikC_1 \exp(-ikL/2) - ikC_2 \exp(ikL/2) = \kappa D_1^- \exp(-\kappa L/2), \end{cases}$$
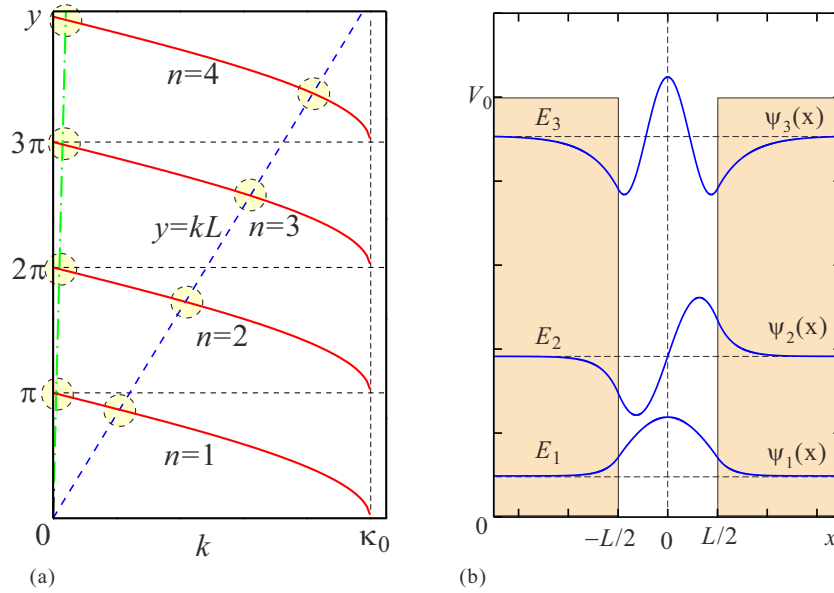
Fig. 7.1 (a) A plot for graphical solutions of k which satisfy eq.(7.4). The crossing points of the functions $-2\arctan(k/\kappa) + n\pi$ and kL give the solutions of (7.4). (b) Bound eigenstates for $n = 1, 2, 3$ under the condition $l = 8$. The baselines for the wavefunctions are the eigenenergies $E_{1,2,3}$ measured with V_0 (for $l = 8$ there are only three bound state solutions, which is different from the situation in the left figure).

respectively. Erasing the constants the following condition is obtained.

$$\exp(2ikL) = \left(\frac{\kappa - ik}{\kappa + ik}\right)^2 = \exp\left(-4i\arctan\frac{k}{\kappa}\right),$$

$$\therefore kL = -2\arctan\frac{k}{\sqrt{\kappa_0^2 - k^2}} + n\pi, \quad \kappa_0^2 \equiv \frac{2mV_0}{\hbar^2}, \quad n = 1, 2, \cdots. \tag{7.4}$$

Let us take kL as a positive value without loosing generality because the solutions contain $-k$ equivalently, and we restrict the value of $\arctan(x)$ between 0 and $\pi/2$. As shown in Fig. 7.1(a), the crossing points of the curves and the line, $-2\arctan(k/\sqrt{\kappa_0^2 - k^2}) + n\pi$ and kL give the values of k, which satisfy (7.4). As easily guessed from the analogy with the case of infinite barriers, even numbers of n correspond to odd parity wavefunctions, while odd numbers correspond to even parities.

In Fig. 7.1(b), we show the form of wavefunctions for the bound states in the case of $l = 8$.

7.1.2 Optical absorption in quantum wells

We would like to have a short look at optical absorption in quantum wells. As usual we take z-axis vertical to the well plane. We write the envelope functions for electrons and holes as $\phi_e(z)$ and $\phi_h(z)$ respectively and then approximate the total wavefunction as

$$\left.\begin{array}{l}\psi_e(\boldsymbol{r}) = \phi_e(z)\exp(i\boldsymbol{k}_{xy} \cdot \boldsymbol{r}_{xy})u_c(\boldsymbol{r}), \\ \psi_h(\boldsymbol{r}) = \phi_h(z)\exp(i\boldsymbol{k}_{xy} \cdot \boldsymbol{r}_{xy})u_v(\boldsymbol{r}).\end{array}\right\} \tag{7.5}$$

u_c, u_v are lattice periodic parts of the Bloch eigenfunction with $\boldsymbol{k} = 0$. Direct type inter-band optical absorption probabilities are proportional to

$$\langle u_c(\boldsymbol{r})|\boldsymbol{\nabla}|u_v(\boldsymbol{r})\rangle \int_{-\infty}^{\infty} dz\phi_e(z)^*\phi_h(z). \tag{7.6}$$

In the case of infinite height barriers, the envelope functions are written as $\sin(n\pi z/L)$, $\cos(l\pi z/L)$ ($n = 2, 4, \cdots$, $l = 1, 3, \cdots$) and the latter integration over z in (7.6) is finite only between electron envelope function and hole envelope
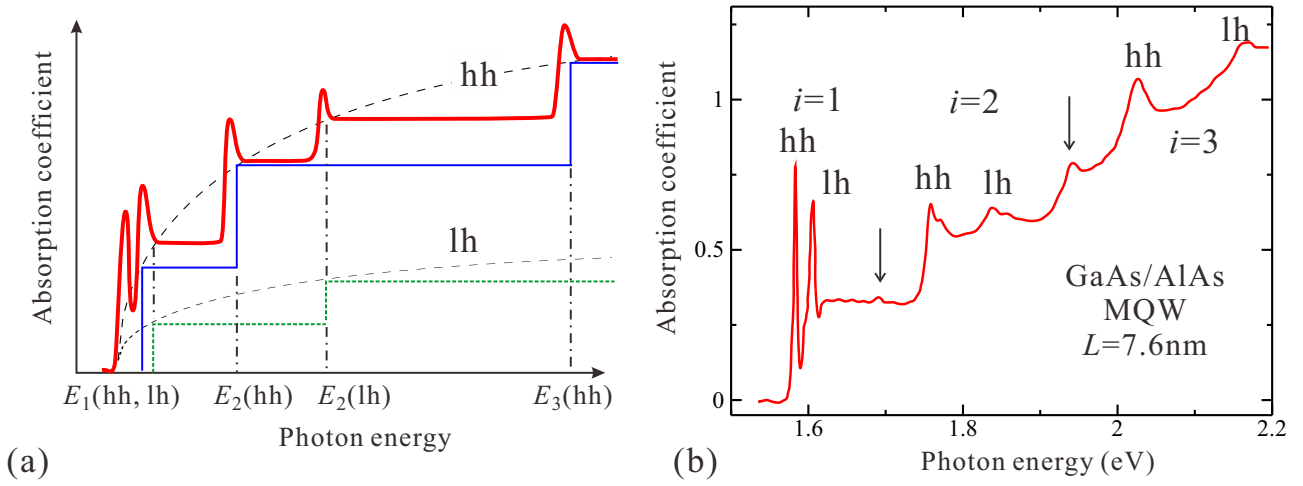
Fig. 7.2 (a) Illustration of theoretically proposed optical absorption spectrum, in which both the coupling density of states and the exciton density of states in the quantum well are taken into account. The approximation that the transition exists only between electrons and holes with the same quantum index. In the valence band of fcc semiconductors we have heavy and ligh holes and transitions with the two bands are considered in the figure. (b) Optical absorption spectrum of a AlAs/GaAs multiple (40 layers) quantum well with width 7.6 nm. The finite barrier height causes transitions between the levels with different quantum indices, which appear in exciton peaks.

function with the same quantum index (n or l in this case). For finite heights, this orthogonality breaks leaving parity selection rule but still elements between different quantum indices are small and we only consider the transition between the states with the same index. The energy associated with the transition is

$$E = E_{\mathrm{g}} + \Delta E_n^{(eh)} + \frac{\hbar^2}{2\mu}k_{xy}^2, \tag{7.7}$$

where $\Delta E_n^{(eh)}$ is the sum of the energies for electron and hole in n-th energy levels, $1/\mu = 1/m_e^* + 1/m_h^*$ is the reduced mass. The last term for two-dimensional kinetic energy indicates that there should be continuous absorption spectrum above $\Delta E_n^{(eh)}$ corresponding to the two-dimensional density of states.

From $E = (\hbar^2/2m^*)k^2$ and $n = \pi k^2/(2\pi)^2 = (E/4\pi)(2m^*/\hbar^2)$, the two-dimensional density of states can be written as

$$\frac{dn}{dE} = \frac{m^*}{2\pi\hbar^2}H(E) \quad (H(x) : \text{Heaviside function}). \tag{7.8}$$

This is constant for energy and with (7.7), we expect a staircase like optical absorption spectrum.

Formaion of excitons appears in optical absorption as peaks at energies lower than the fundamental absorption edge. Such peaks for excitons in quantum wells are illustrated in Fig. 7.2(a). Only the ground states ($n = 0$) of the excitons are considered. And coupling density of states between electrons and holes with different subband quantum indices is ignored assuming that the barrier is high enough. Figure 7.2(b) shows an experimental result on an AlGaAs/GaAs multiple quantum well with width 7.6 nm. The lineshape of the absorption spectrum can be understood as an overlap of staircase-like shape reflecting the two-dimensional density of states (7.8) and absorption by excitons indicated as hh or lh. Because the barrier height is finite in the experiment, peaks due to the transition between states with different quantum indices are also observed. The effect of low-dimensionality is observable in increases of binding energy of excitons, which results in wider separation of exciton peaks from absorption edges and the peaks persist up to higher temperatures.

Now we can see that the optical absorption spectra can provide experimental determination of band-discontinuities ΔE_{c}, ΔE_{v}. In the combination of GaAs-Al$_x$Ga$_{1-x}$As, reseachers could not separate lh and hh peaks in very early experiments presumably due to low quality of samples. The result once led them to a wrong conclusion of $\Delta E_{\mathrm{c}} : \Delta E_{\mathrm{v}} = 85 : 15$ because ΔE_{v} should be too small to accomodate the lh level. After the revised experiments, it was established

that $\Delta E_c : \Delta E_v = 57 : 43$ is a good empirical law.

Appendix 6C: Trials for simple theory to find band discontinuities

The capacity of computers has increased dramatically, and first-principles calculations that require a lot of computer resources as LAPW, can now be performed relatively easily. Even though, the researchers are still trying to construct a theory to obtain band discontinuity from a small number of experimental parameters using simple physical principles. I will introduce such researches so far, but it has been found that many cannot withstand the subsequent criticisms, experiments, and first-principles calculations.

6C.1 Common anion rule

This "common anion rule" is considered for compound semiconductors[8] which have finite inonicity. The claim is as follows. Because the valence band is mostly composed of p-orbitals of anion, $\Delta E_v \approx 0$ for the semiconductors with a common anion. This is a surprizingly rough theory. The prediction is far from experiments and from other models.

6C.2 Pseudo-potential theory

The quantitiy "affinity" can be formally calculated from the first principles. In the era of Anderson's research, ΔE_c are determined from the experiments, and the affinity is obtained from the fitting. Here we see a theory, which is constructed by Frensley and Krömer[9] aiming at finding band discontinuity from bulk parameters.

The calculation goes as follows. First with the self-consistent pseudo-potential method, the relative positions of bulk bands are calculated in the electrostatic potential inside the crystal[10]. Next from the electronegativity of consisting atoms and from the band structure, the electrostatic potential at the interface is calculated and the relative positions of bands are obtained. They claimed the agreement with exeriments[9].

6C.3 LCAO theory

W. A. Harrison applied his linear combination of atomic orbitals (LCAO) theory to the heterostructure in ref.[11, 12]. In Harrison's theory, LCAO forms the bands. Most of bands in semiconductors can be expressed with the combination of the single s-orbital and the three p-orbitals. The valence band top is composed of p-orbitals and expressed as

$$E_v = \frac{\epsilon_p^c + \epsilon_p^a}{2} - \left[\left(\frac{\epsilon_p^c - \epsilon_p^a}{2} \right)^2 + V_{xx}^2 \right]^{1/2}, \tag{6C.1}$$

where $\epsilon_p^{c,a}$ are the energies of p-orbitals of cation and anion respectively on their own sites, V_{xx} is the matrix element between neighboring p-orbitals. According to the theory[11], V_{xx} is approximated as

$$V_{xx} = 2.16\hbar^2/md^2, \tag{6C.2}$$

where m is the electron mass and d is the bond length. The number 2.16 is obtained by fitting the results for Si and Ge to those of other band calculation[13].

In this method, the valence band discontinuities for many semiconductor can be easily calculated and often used for the estimation. It is said to give fairly good agreements with experiments in many cases though with many exceptions *e.g.*, in the case of GaAs-AlAs the Harrison theory gives ΔE_v=0.04 eV for about 0.5 eV in the experiments[14]. Actually the main difference in ref. (6C.1) from the common anion rule is just ϵ_p^c and still the approximation is rough.

6C.4   Interface dipole theory

This theory had been pur forward by Tersoff and Harrison[15, 16, 17]. First Tersoff criticized Harrison's LCAO theory that the theory is not reallistic in that the charge transfer between the semiconductors is ignored and no dipole exists at the interface[15]. Then they collaborated in constructing LCAO theories with electric dipoles[17].

In the Tersoff's original idea, on the surface of a semiconductor (an insulator), an energy level of "charge neutrality" can be defined in the band gap. The charge neutrality level is given at the point where the contributions of the valence orbitals and those of the conduction orbitals are balanced. In a metal-semiconductor Schottky junction, this charge neutrality level should be matched to the Fermi surface in the metal[*3]. The charge transfer through the interface is over a very short distance, only single lattice constant. The transfer length scale is very different from the space-charge created so as to match E_F with the bulk value. When two semiconductors are jointed, there is no charge transfer for matched charge neutrality levels. Otherwise the transfer occurs to match the charge neutrality levels. Hence, if we can calculate the position of charge neutrality level, the band offset can be derived from that.

	E_B	$E_F(\mathrm{Au})$[a]	$E_F(\mathrm{Al})$[a]
Si	0.36	0.32	0.40
Ge	0.18	0.07	0.18
AlAs	1.05	0.96	
GaAs	0.70	0.52	0.62
InAs	0.50	0.47	
GaSb	0.07	0.07	
GaP	0.81	0.94	1.17
InP	0.76	0.77	

[a]Reference 18.

	Experiments	Theory	Difference
AlAs/GaAs	0.19[b]	0.35	0.16
InAs/GaSb	0.51	0.43	−0.08
GaAs/InAs	0.17	0.20	0.03
Si/Ge	0.20	0.18	−0.02
GaAs/Ge	0.53	0.52	−0.01

[a]Reference 1.
[b]However, see text and Refs. 1, 19, and 20.

(a)   (b)

Tab.   6C.1   (a) The in-gap state E_B obtained by equating the contributions from the both bands to (6C.3) and the positions of E_F measured in the Schottky diodes with Au and Al as the electrodes.

For the calculation of the charge neutrality level (or, metal-induced gap states, MIGS), the contributions from the valence and the conduction to the real-space averaged Green function

$$G(\boldsymbol{R}, E) = \int d^3r \sum_{n\boldsymbol{k}} \frac{\psi_{n\boldsymbol{k}}^*(\boldsymbol{r})\psi_{n\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{R})}{E - E_{n\boldsymbol{k}}} = \sum_{n\boldsymbol{k}} \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{R}}}{E - E_{n\boldsymbol{k}}} \tag{6C.3}$$

are equalized to give MIGS E_B. E_B obtained with this method and the positions of E_F obtained in Schottky diodes with Au and Al electrodes[?] are listed in Tab. 6C.1(a). Already in this table, the agreement is not very good. And after the publication, there occured many criticisms including the experiments. In conclusion, the theory is convenient in the discussion of chemical trend but it is hard to say that it can be used for device design.

6C.5   Example of first principles calculation

Wei and Zunger proceeded with the so-called first-principles calculation of the interface, and the low accuracy of the common anion law and even the simple LCAO theory is due to the roughness of the bulk band calculation rather than the effect of the interface dipole[19]. That is, the bulk contribution $\Delta E_\mathrm{VBM}^\mathrm{b}$ and the surface contribution $\Delta E_\mathrm{VBM}^\mathrm{is}$ to the energy difference ΔE_VBM at the valence band maximum (VBM) are in the relation

$$\Delta E_\mathrm{VBM} = \Delta E_\mathrm{VBM}^\mathrm{b} + \Delta E_\mathrm{VBM}^\mathrm{is}. \tag{6C.4}$$

[*3] In the most of real Schottky junctions, there are defect levels with very high densities and the Fermi levels are pinned there. At the heterointerface with small defect densities, the situation is different.

According to their claim, $\Delta E_{\text{VBM}}^{\text{is}}$ is small and the problem in LCAO theory rather lies in the estimation of $\Delta E_{\text{VBM}}^{\text{b}}$. That is in Harrison theory, only s and p orbitals are considered but particularly the contribution from the d orbitals of cation is comparatively large and the most of disagreement with experiments can be explained with this (calculation was done by all-electron generalized linear augmented plane wave method[20]).

Systems	Tight-binding[a]			$\Delta E_{\text{VMB}}^{\text{expt}}$	Average (with SO)	All-electron (Present results)					δ_{pd}
	$\Delta E_{\text{VBM}}^{b}$	$\Delta E_{\text{VMB}}^{\text{IS}}$	$\Delta E_{\text{VMB}}^{\text{tot}}$			Average (no SO)	Using 1s	Using 2s	Using 3$p_{1/2}$		
CdTe-HgTe	0.00	0.09	0.09	$0.35 \pm 0.06^{\text{b}}$	0.37	0.39	0.377	0.388	0.400	0.34
CdTe-ZnTe	−0.07	0.00	−0.07	$\cdots$	0.13	0.12	0.125	0.122	0.108	0.04
ZnTe-HgTe	0.07	0.09	0.16	$\cdots$	0.26	0.29	0.277	0.286	0.289	0.30
AlAs-GaAs	0.01	0.15	0.16	$0.45 \pm 0.05^{\text{c}}$	0.42	0.41	0.41	0.40	$\cdots$	0.31

Tab. 6C.2 ΔE_{v} for semiconductors with Te and As as anion. The results of simple LCAO, experiments, and all-electron first principles calculation.

The calculated results are summarized in Tab. 6C.2. Now the LAPW method can be rather easily utilized in the form of convenient packages like HiLAPW or VASP though still consumes large calculation resources and the jobs are heavy). The method is only for periodic systems and in the case of heterointerface, the unit cell is taken large along vertical direction to the interfaceso as to contain two intefaces and the periodic boundary condition is applied. This is, in a sense, calculation of a superlattice band structure and can be used to check the staircase approximation of the heterointerface.

Appendix 6D: Recombination current and ideality factor

In the discussion of current-voltage characteristics of pn junctions in the text, we only considered the diffusion current. In reallistic pn-junctions, various other factors contribut the current. Here we have a brief look at the current caused by carrier recombination in the depletion layer at the junction interfaces.

First we consider direct gap semiconductors, in which the interband recombination rate is much higher than those in indirect ones. Let the interband recombination rate be R_e, this should be proportional to the carrier concentrations n and p. Thus R_e is proportional to the product pn. Let R_{rc} be the coefficient, then

$$R_e = R_{\text{rc}} pn. \tag{6D.1}$$

R_e equals to the thermal activation rate G_{th} of the electron-hole pair in the dark and in equilibrium. Then the law of mass action gives

$$R_{\text{rc}} = \frac{G_{\text{th}}}{pn} = \frac{G_{\text{th}}}{n_i^2}. \tag{6D.2}$$

When there is optical activation or minority carrier injection by the external current, the activation rate and the recombination rate are not balanced and the difference is the net recombination rate U. In n-type semiconductors, the variation in the hole concentration is the main factor. If we write $p_n = p_0 + \Delta p$, $n_n \approx N_{\text{D}}$, then

$$U = R_e - G_{\text{th}} = R_{\text{rc}}(pn - n_i^2) \approx R_{\text{rc}} \Delta p N_{\text{D}} \equiv \frac{\Delta p}{\tau_p}, \tag{6D.3}$$

where we define the minority carrier lifetime as

$$\tau_p = \frac{1}{R_{\text{rc}} N_{\text{D}}}. \tag{6D.4}$$

Similarly the electron lifetime in p-type semiconductors are written as

$$\tau_n = \frac{1}{R_{\text{rc}} N_{\text{A}}}. \tag{6D.5}$$

In contrast, in the indirect gap semiconductors like Si or Ge, the carrier recombination is via the localized traps. In that case, the net recombinatin rate is, according to so called Shockley-Read-Hall statistics[21] written as

$$U = \frac{\sigma_n \sigma_p v_{\text{th}} N_t (pn - n_i^2)}{\sigma_n \left[n + n_i \exp \dfrac{E_t - E_i}{k_B T} \right] + \sigma_p \left[p + n_i \exp \dfrac{E_i - E_t}{k_B T} \right]}, \tag{6D.6}$$

where N_t is the trap density, σ_n, σ_p are the capture cross-sections for electrons and holes respectively, E_t is the trap level, E_i is the Femi level of the intrinsic semiconductors. And v_{the} is the thermal velocity of the minority carrier

$$v_{\text{th}} = \sqrt{\frac{3 k_B T}{m^*}}. \tag{6D.7}$$

In eq.(6D.6), U takes the maximum at $E_t \approx E_i$. Though actually E_t distribute over the band gap, the trap levels close to E_i contribute lartgely to U [*4], then as a coarse approximation, we consider only single species of traps and put $E_t = E_i$ then

$$U = \frac{\sigma_n \sigma_p v_{\text{th}} N_t (pn - n_i^2)}{\sigma_n (n + n_i) + \sigma_p (p + n_i)}. \tag{6D.8}$$

Just like in the case of interband transition, we write U as $\Delta p / \tau_p$ or $\Delta n / \tau_n$, giving

$$\tau_p = \frac{1}{\sigma_p v_{\text{th}} N_t}, \quad \tau_n = \frac{1}{\sigma_n v_{\text{th}} N_t}. \tag{6D.9}$$

Now we use quasi-Fermi levels introduced in eq.(6.4) and from eq.(3.13) the np product is written as

$$np = n_i^2 \exp \frac{\mu_e - \mu_h}{k_B T}. \tag{6D.10}$$

Substitutin the above into (6D.6), we obtain

$$U = \frac{\sigma_n \sigma_p v_{\text{th}} N_t n_i^2 \left[\exp \dfrac{eV}{k_B T} - 1 \right]}{\sigma_n \left[n + n_i \exp \dfrac{E_t - E_i}{k_B T} \right] + \sigma_p \left[p + n_i \exp \dfrac{E_i - E_t}{k_B T} \right]}. \tag{6D.11}$$

Then again we put $E_t = E_i$, and for further simplicity, we assume $\sigma_n = \sigma_p = \sigma$ to obtain

$$U = \frac{\sigma v_{\text{th}} N_t n_i^2 \left[\exp \dfrac{eV}{k_B T} - 1 \right]}{n + p + 2 n_i} = \frac{\sigma v_{\text{th}} N_t n_i^2 \left[\exp \dfrac{eV}{k_B T} - 1 \right]}{n_i \left[\exp \dfrac{\mu_e - E_i}{k_B T} + \exp \dfrac{E_i - \mu_h}{k_B T} + 2 \right]}. \tag{6D.12}$$

Further, when μ_e, mu_h are position dependent, U takes the maximum in the case E_i places in the middle between μ_e and μ_h. Then (6D.12) reduces to

$$U \approx \frac{\sigma v_{\text{th}} N_t n_i^2 \left[\exp \dfrac{eV}{k_B T} - 1 \right]}{2 n_i \left[\exp \dfrac{eV}{2 k_B T} + 2 \right]} \approx \frac{1}{2} \sigma v_{\text{th}} N_t n_i \exp \frac{eV}{2 k_B T} \quad eV > k_B T. \tag{6D.13}$$

Because the above is the maximum the estimation should be a bit large but the current density due to the recombination can be written as

$$j_{\text{rc}} = \int_0^{w_d} q U dx \approx \frac{q w_d n_i}{2 \tau} \exp \frac{eV}{2 k_B T}. \tag{6D.14}$$

In eq. (6D.14) in comparison with eq. (6.11), the voltage term in the exponential has an extra factor 1/2. To put it plainly, this is because the energy exchange when recombination occurs in the trap is half that of the case where the

---

[*4] This means that the lifetime of minority carriers is determined by a deep level, especially in indirect semiconductors. In "solar grade" Si, in conparison with those for LSI (impurity $^{-10} \sim 10^{-11}$), the purity can be a bit lower while the deep level concentrations should be reduced.

current is generated by overcoming the bandgap (6.11). In this way, different processes in which current flows generally have different voltage coefficients. Then in experiments, the forward current is written as

$$J_{\mathrm{F}} \propto \exp \frac{eV}{\eta k_{\mathrm{B}} T} \tag{6D.15}$$

and the factor η (**ideality factor**) is fit to the experiment. When η is close to 1, the diffusion current is dominant and the junction is close to the ideal case. When it is close to 2, the recombination current inside the depletion layer is dominant.

In the laboratories, η sometimes goes over 2 and still takes higher values. In the case of pn-junctions, the interface comes to the middle of depletion layer and there is some interdiffusion of dopants, thus the factor 2 is frequently obtained while it is usually close to 1 in the case of Schottky junctions.

References

[1] 勝本信吾「半導体量子輸送物性」(培風館，2014).

[2] G. Bastard, "Wave Mechanics Applied to Semiconductor Heterostructures" (Editions de Physique, France, 1990).

[3] R. L. Anderson, IBM J. Res. Dev. **4**, 283 (1960); Solid-State Electronics **5**, 341 (1962).

[4] C. Liu *et al.*, Phys. Rev. Lett. **100**, 236601 (2008).

[5] J.E. Matthews, A.E. Blakeslee, J. Crystal Growth **27**, 118 (1974).

[6] H. C. Casey, Jr., M. B. Panish, "Heterostructure Lasers" Part B (Academic Press, 1978).

[7] A. K. Geim and I. V. Grigorieva, Nature **499**, 419 (2013).

[8] A. G. Milnes and D. L. Feucht, "Heterojunctions and Metal Semiconductor Junctions" (Academic Press, 2012).

[9] W. R. Frensley and H. Kroemer, Phys. Rev. B **16**, 2642 (1977).

[10] M. L. Cohen and J. R. Chelikowsky, "Electronic Structure and Optical Properties of Semiconductors" (Springer, 1989).

[11] W. A. Harrison, J. Vac. Sci. Tech. **14**, 1016 (1977).

[12] W. A. Harrison *et al.*, Phys. Rev. B **18**, 4402 (1978).

[13] D. J. Chadi and M. L. Cohen, phys. stat. solidi (b) **68**, 405 (1975).

[14] 竹田 美和, 応用物理 **67**, 1077 (1998).

[15] J. Tersoff, Phys. Rev. B **30**, 4874 (1984).

[16] J. Tersoff, Phys. Rev. Lett. **56**, 2755 (1986).

[17] W. A. Harrison and J. Tersoff, J. Vac. Sci. Tech. B **4**, 1068 (1986).

[18] S. M. Sze and K. K. Ng, "Physics of Semiconductor Devices" 3rd ed. (Wiley, 2008).

[19] S.-H. Wei and A. Zunger, Phys. Rev. Lett. **59**, 144 (1987).

[20] S.-H. Wei, H. Krakauer, and M. Weinert, Phys. Rev. B **32**, 7792 (1985).

[21] W. Schockley and W. T. Read, Phys. Rev. **87**, 835 (1952); R. N. Hall, Phys. Rev. **87**, 387 (1952).