Physics of Semiconductors (9)

Shingo Katsumoto
Dept. Physics and Inst. for Solid State Physics, University of Tokyo

July 17, 2016

Ch.3 Heterojunction and quantum confinement to two-dimensional systems

Today, I would like to introduce how to produce two-dimensional systems of electrons or holes with emphasis on heterojunction. Porf. Akiyama gave an explanation on the treatment of heterojunctions based on k·p perturbation and I will not revisit it though I would like to confirm what we are treating when we see heterojunctions as simple step function potential. In very short, in a Bloch-type wavefunction we drop the lattice periodic functions renormalizing their effect into "band structure" and the residual plane waves are "envelope functions". Then we can map the problem of heterojunction into a potential problem of envelope functions. As noted by Prof. Akiyama, connection of envelope functions at the interface contains subtle problems and we should pay attention that envelope functions are not wavefunction themselves.

3.1 Heterojunction

3.1.1 Envelope function (review)

Envelope function is one of the basic concepts in treating heterojunctions. Prof. Akiyama already gave a detailed explanation on the k·p treatment of heterojunctions and here we only review the concept. Let $u_{n\boldsymbol{k}}(\boldsymbol{r})\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ ($u_{n\boldsymbol{k}}(\boldsymbol{r})$: lattice periodic function) be a Bloch-type lattice eigen function. n is the band index. Around the band edge, $u_{n\boldsymbol{k}}(\boldsymbol{r})$ mainly reflects lattice-specific band structures and the plane wave part represents spatial modulation of the band-edge properties. When there is a spatial modulation in material parameters, $u_{n\boldsymbol{k}}(\boldsymbol{r})$ gets the modulation and the effect is absorbed as if the modulation generates a potential for the plane wave part. Now if we forget the lattice periodic function and think the "plane wave" part as a "wavefunction" we can simplify the problem to a potential problem of quantum mechanics. The "plane wave part" is the envelope function to be exact. Band calculations for heterojunctions are surprizingly difficult but once it is renormalized into a potential problem, the rest is just elementary quantum mechanics.

Let us review. Consider a perturbation potential $U(\boldsymbol{r})$ is added to the lattice Hamiltonian $\mathscr{H}_0$. Let $\eta(\boldsymbol{r})$ be an eigenstate of the system and

$$[\mathscr{H}_0 + U(\boldsymbol{r})]\eta(\boldsymbol{r}) = E\eta(\boldsymbol{r}). \tag{3.1}$$

$\eta(\boldsymbol{r})$ can be expanded with $|n,\boldsymbol{k}\rangle$ because of their completeness as

$$\eta(\boldsymbol{r}) = \sum_{n,\boldsymbol{k}} f(n,\boldsymbol{k})|n,\boldsymbol{k}\rangle. \tag{3.2}$$

Substituting the above into (3.1) and operation of $\langle n',\boldsymbol{k}'|$ give the next expression:

$$[E_0(n',\boldsymbol{k}') - E]f(n',\boldsymbol{k}') + \sum_{n,\boldsymbol{k}}\langle n',\boldsymbol{k}'|U|n,\boldsymbol{k}\rangle f(n,\boldsymbol{k}) = 0. \tag{3.3}$$

We replace $u_{n\boldsymbol{k}}$ with $u_{n0}(\boldsymbol{r})$ assuming weak dependence of $\boldsymbol{k}$ on $\boldsymbol{k}$ and ignore interband mixing to obtain

$$\eta_n(\boldsymbol{r}) = u_{n0}\sum_{\boldsymbol{k}} f(n,\boldsymbol{k})e^{i\boldsymbol{k}\boldsymbol{r}} \equiv u_{n0}f_n(\boldsymbol{r}), \tag{3.4}$$

where $f_n(\boldsymbol{r})$ is called **envelope function**. Envelope function is a summation of plane wave parts in the Bloch functions for including the effect of $U(\boldsymbol{r})$ (formally inverse Fourier transform) From (3.3), dropping n', we obtain

$$\frac{\hbar^2 k'^2}{2m^*} f(\boldsymbol{k}') + \sum_{\boldsymbol{k}} U_{\boldsymbol{k}'-\boldsymbol{k}} f(\boldsymbol{k}) = E f(\boldsymbol{k}'). \tag{3.5}$$

Inverse Fourier transform of the both sides on $\boldsymbol{k}'$ gives

$$\left[-\frac{\hbar^2 \nabla^2}{2m^*} + U(\boldsymbol{r}) \right] f(\boldsymbol{r}) = E f(\boldsymbol{r}). \tag{3.6}$$

Note that the second term in (3.5) is a convolution. Equation (3.6) is nothing but the Schrödinger equation for a particle with the effective mass in the potential $U(\boldsymbol{r})$. The renormalization of the problem into envelope function is hence called **effective mass approximation**.

3.1.2 Anderson's rule for heterojunction

Equation (3.6) means when there is some spatial structure other than the lattice potential of matrix material, we can treat it as a potential problem of a particle with the effective mass. If the potential is just flat or very soft, the envelope function is almost a plane wave or a free particle. When we considered pn-junctions we adopted a rigid band model [1], in which the energy band has spatial slope keeping the band gaps. This approximation is based on the effective mass approximation.

Prof. Akiyama also gave an explanation on whether this approximation is applicable or not to the heterojunction problem and if it is, how. Generally there is an abrupt variation in the lattice periodic function $u_{n\boldsymbol{k}}(\boldsymbol{r})$ and direct application is questionable. Even if we succeed in renormalizing the effect into a step function potential, that should have a discontinuous change and requires summation over a wide range of $\boldsymbol{k}$ to obtain an envelope function and replacement of $u_{n\boldsymbol{k}}$ with u_{n0} also becomes a problem. I'll skip the discussion but to summarise, whether we can apply the effective mass approximation to a heterojunction depends on the combination of materials which compose the junction. In the case of GaAs/$Al_xGa_{1-x}As$, it is established that the step function potential and continuous envelope function and the derivative over the junction interface is a good approximation. In general combination, the simplest approximation that continuous and differentiable envelope function does not hold but with some additional rule, similar treatment is available.

In summary, the model in which band gaps and effective masses are specific to materials and "rigid" to interfaces, is mostly adopted in treatments of heterojunctions. In real many electron systems, we consider envelope functions (effective mass approximation) as one-electron functions to form Slater determinant.

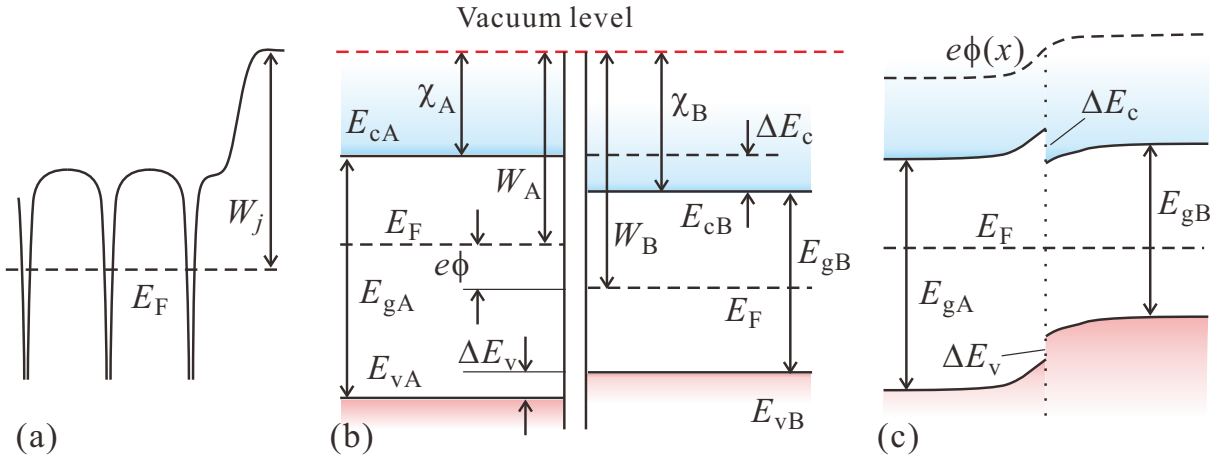


Figure 3.1: A simple model of heterojunction (Anderson's rule). (a) Energies required for extraction of an electron from the Fermi energy to the vacuum level (workfunction W_j). (b) Band alignment assuming fixed electron affinity $\chi_{A,B}$. (c) In equilibrium, E_F should be constant throughout the junction and accordingly junction potential, in other words band bending appears.

[1] In solid state physics, the term "rigid band model" is mostly used in the mixing of energy of bands in the formation of alloys. The present usage is rather minor.

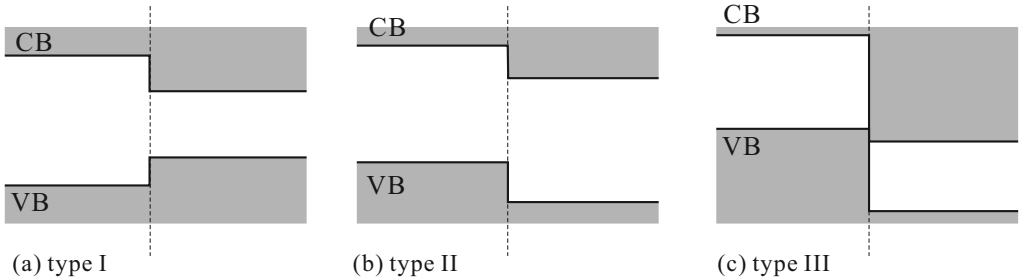


Figure 3.2: Three types of heterojunctions. (a) type I. Ordinary alignment. The potential is lower in the narrower side both for electrons (conduction band) and holes (valence band). (b) type II. Staggered alignment. At one side, the potential for electrons is lower while that for holes is higher. (c) type III. Gap disappearance (semi-metal). The conduction band in one side is connected to the valence band in the other side. The bandgap disappears consequently.

What we should consider next, in this model, is to find appropriate values for $\Delta E_{\rm c}$ and $\Delta E_{\rm v}$ for a given combination of semiconductors. Knowing band gaps $E_{\rm gA}$ and $E_{\rm gB}$ respectively for semiconductors A and B, we put

$$E_{\rm gB} - E_{\rm gA} = \Delta E_{\rm c} + \Delta E_{\rm v}.$$

Therefore another condition gives the values. An attempt to give the condition from experimental data phenomenologically was made by R. L. Anderson in the age they did not know the band structure even in the bulk. The method is illustration in Fig.3.1and called Anderson's rule. To extract an electron at the conduction ban bottom into vacuum needs electron affinity $\chi_{\rm A}$ for A, $\chi_{\rm B}$ for B. These quantities can be measured *e.g.* with photoelectron spectroscopy. Then assuming the vacuum level is common for A and B,

$$\Delta E_{\rm c} = \chi_{\rm B} - \chi_{\rm A}.$$

With the positions of $E_{\rm F}$, we can draw the band diagram as in Fig.3.1(c) from the condition that $E_{\rm F}$ should be flat throughout the junction.

Many experimental results support the step function potential model with fixed $\Delta E_{\rm c}$, $\Delta E_{\rm v}$ sometimes with additional rules for envelope function connection. However, appropriate values for $\Delta E_{\rm c}$ etc. found in experiments are significantly shifted from Anderson's rule. Following Anderson's rule, various conventional methods like common anion rule, surface states rule etc. have been proposed though still there is no easy-going method with high accuracy for interfaces, which break the translational symmetry. That is we still need to adopt large scale first principle calculations to obtain reliable values for band discontinuities. Later we will see some examples of the determination in experiments.

3.1.3 Types of heterojunctions

Band alignment of heterojunctions can be classified into three types illustrated in Fig.3.2. One of the representative combinations GaAs-$Al_xGa_{1-x}As$ has the alignment of type-I (Fig.3.2(a)). A GaAs layer sandwiched by two $Al_xGa_{1-x}As$ layers thus works as a quantum well both for electrons and holes. On the other hand, GaSb and narrow gap InAs form a junction of type-III in Fig.3.2(c), in which the conduction band of InAs and the valence band of GaSb overlap and the energy gap is "shorted" at the interface. If we adopt a mixed crystal $Al_xGa_{1-x}Sb$ instead of GaSb and make the gap larger, the valence band top lowers and at $x =$**, the top crosses that in the InAs and type-II interface is realized.

3.2 Formation of heterojunctions

3.2.1 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.

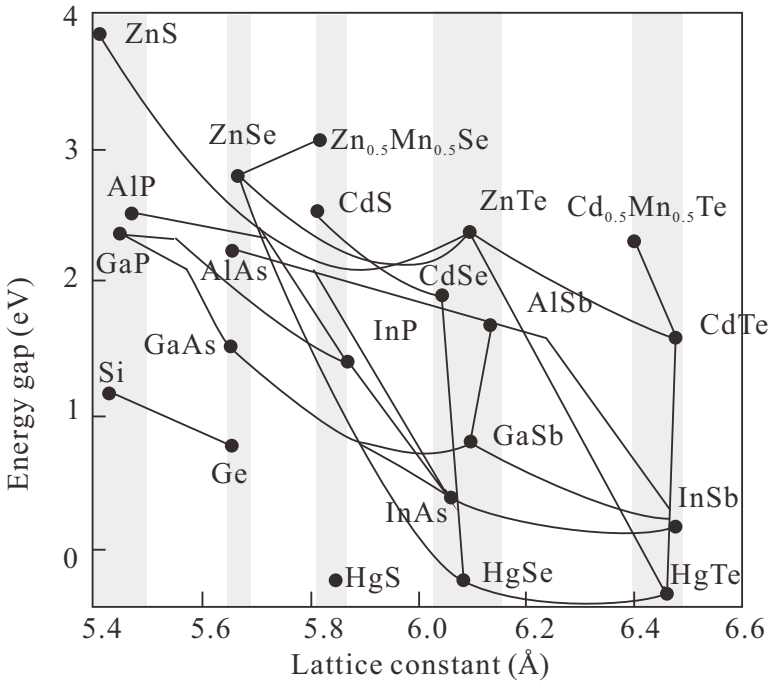


Figure 3.3: Plots of the lattice constants and the energy gaps of II-VI, III-V compound semiconductors and IV elemental semiconductors. The lines connecting the points indicate possible mixed crystals. Vertical gray bands indicate possible groups of lattice matched heterostructure growth.

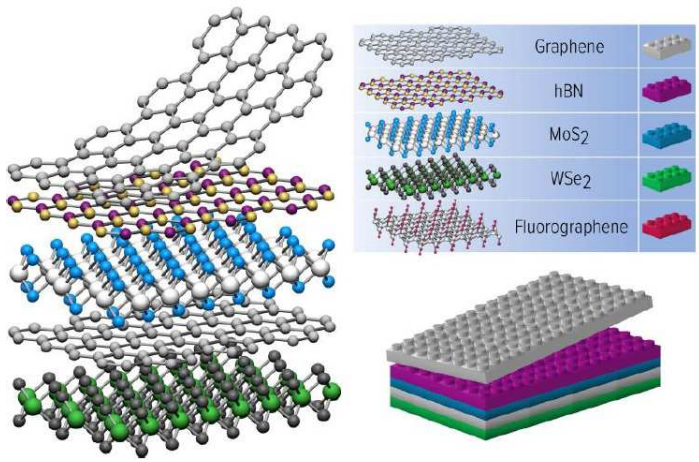


Figure 3.4: Conceptual illustration of van der Waals heterostructure, which is produced by stacking various two-dimensional materials.

An important point in the formation of heterojunction is the **lattice matching** in lattice constants and crystal systems. In Fig.3.3, 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**[2]. 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 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[3].

3.2.2 van der Waals heterostructure

Recently van der Waals heterostructure, which is formed in completely different way, is collecting attentions[4]. That is a mechanical stacking of two-dimensional materials like graphene as shown in Fig.3.4 (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.

3.3 Qauntum 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.

3.3.1 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

$$\left[-\frac{\hbar^2 \mathrm{d}^2}{2m\mathrm{d}x^2} + V_0\right]\psi = E\psi. \tag{3.7}$$

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

$$\psi(x) = \begin{cases} C_1\exp(i\kappa x) + C_2\exp(-i\kappa x) & E \geq V_0, \\ D_1\exp(\kappa x) + D_2\exp(-\kappa x) & E < V_0. \end{cases} \tag{3.8}$$

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

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

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

$$\psi = C_1\exp(ikx) + C_2\exp(-ikx), \quad k \equiv \frac{\sqrt{2mE}}{\hbar}, \tag{3.9}$$

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

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

$$\text{Differentiability} \begin{cases} ikC_1\exp(ikL/2) - ikC_2\exp(-ikL/2) = -\kappa D_2^+\exp(-\kappa L/2), \\ ikC_1\exp(-ikL/2) - ikC_2\exp(ikL/2) = \kappa D_1^-\exp(-\kappa L/2), \end{cases}$$

respectively. Erasing the constants the following condition is obtained.

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

$$\therefore kL = -2\arctan\frac{k}{\sqrt{\kappa_0^2 - k^2}} + n\pi, \qquad \kappa_0^2 \equiv \frac{2mV_0}{\hbar^2}, \quad n = 1, 2, \cdots. \tag{3.10}$$

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.3.5(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 (3.10). 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.3.5(b), we show the form of wavefunctions for the bound states in the case of $l = 8$.

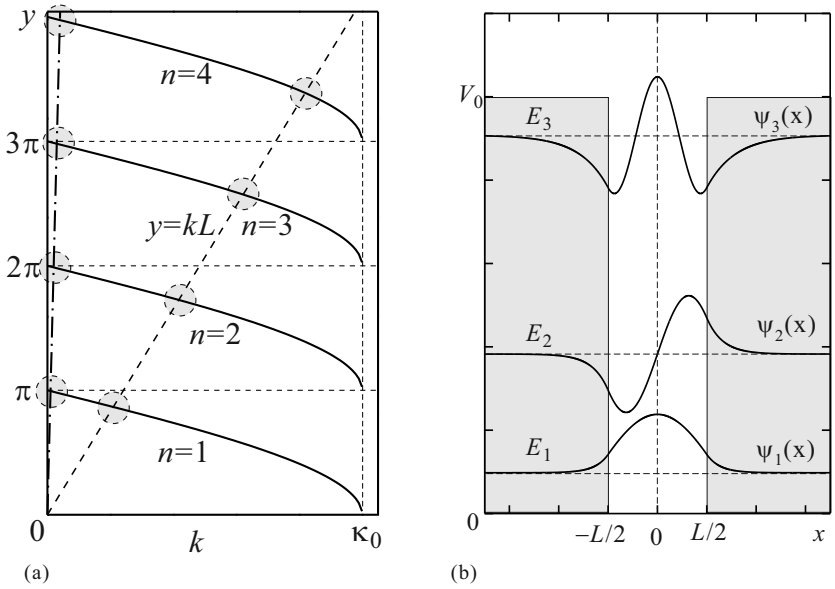


Figure 3.5: (a) A plot for graphical solutions of k which satisfy eq.(3.10). The crossing points of the functions $-2\arctan(k/\kappa) + n\pi$ and kL give the solutions of (3.10). (b) Bound eigenstates for $n = 1, 2, 3$ under the condition $l = 8$. The baselines for the wavefunctions are the eigenenergies $E_{1,2,3}$ measured with V_0 (for $l = 8$ there are only three bound state solutions, which is different from the situation in the left figure).

3.3.2 Optical absorption in quantum wells

In spite of the principle of "not going into opto-material science" in this lecture, 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

$$\left.\begin{aligned} \psi_e(\boldsymbol{r}) &= \phi_e(z)\exp(i\boldsymbol{k}_{xy}\cdot\boldsymbol{r}_{xy})u_c(\boldsymbol{r}), \\ \psi_h(\boldsymbol{r}) &= \phi_h(z)\exp(i\boldsymbol{k}_{xy}\cdot\boldsymbol{r}_{xy})u_v(\boldsymbol{r}). \end{aligned}\right\} \tag{3.11}$$

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

$$\langle u_c(\boldsymbol{r})|\boldsymbol{\nabla}|u_v(\boldsymbol{r})\rangle \int_{-\infty}^{\infty} dz\phi_e(z)^*\phi_h(z). \tag{3.12}$$

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 (3.12) is finite only between electron envelope function and hole envelope 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

$$E = E_{\rm g} + \Delta E_n^{(eh)} + \frac{\hbar^2}{2\mu}k_{xy}^2, \tag{3.13}$$

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

$$\frac{dn}{dE} = \frac{m^*}{2\pi\hbar^2}H(E) \quad (H(x) : \text{Heaviside function}). \tag{3.14}$$

This is constant for energy and with (3.13), 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.3.6(a). Only the ground states ($n = 0$) of the excitons

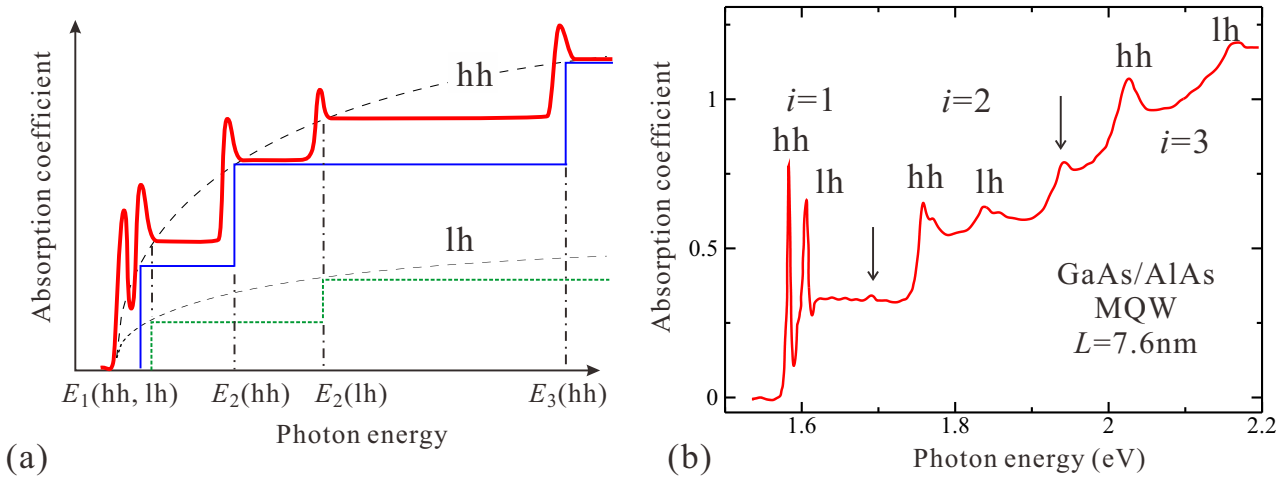


Figure 3.6: (a) Illustration of theoretically proposed optical absorption spectrum, in which both the coupling density of states and the exciton density of states in the quantum well are taken into account. The approximation that the transition exists only between electrons and holes with the same quantum 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.

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 3.6(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 (3.14) 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-$Al_xGa_{1-x}As$, 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 $\Delta E_c : \Delta E_v = 85 : 15$ 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.

3.4 Quantum barrier

"Upside down" of a quantum well potential gives a quantum barrier potential. In the quantum well problem, the focus was on the bound states inside the well while in quantum barriers we see characteristic tunneling phenomena in the upside-down states of **resonant scattering**.

3.4.1 Transfer matrix

Let us consider a region Q in a one-dimensional space and as shown in Fig.3.8(a), and incoming wavefunction $A(k)$ with wavenumber k from the left hand side (LHS), outgoing wavefunction $A_2(k)$ to the right hand side (RHS), and $B_2(k)$, $B_1(k)$ for the other way around. Here we take the momentum k to be common for the momentum conservation. The suffices 1 and 2 indicates the boundaries 1 and 2.

Let us take for an example that a rectangular barrier with width L, and height V_0. Let the wavefunction inside the barrier be $V_i(\kappa) + W_i(\kappa)$. V, W correspond to $e^{-\kappa x}$, $e^{\kappa x}$ respectively and from the Schrödinger equation, $\partial V_i/\partial x = -\kappa V_i$, $\partial W_i/\partial x = \kappa W_i$. The suffix i indicates positions in real space, just as above, putting 1 and 2 to the left and the right edges of the barrier and

$$V_2 = V_1 e^{-\kappa L}, \qquad W_2 = W_1 e^{\kappa L}.$$

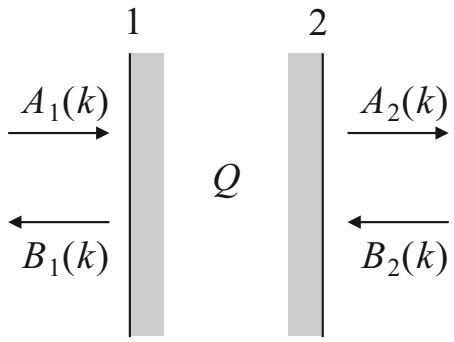


Figure 3.7: Scheme of T-matrix

Now the boundary condition can be written as $\partial A_{1,2}/\partial x = ikA_{1,2}$, $\partial B_{1,2}/\partial x = -ikB_{1,2}$, hence,

$$A_1 + B_1 = V_1 + W_1, \qquad A_2 + B_2 = e^{-\kappa L}V_1 + e^{\kappa L}W_1, \tag{3.15}$$

$$ik(A_1 - B_1) = \kappa(-V_1 - W_1), \qquad ik(A_2 - B_2) = \kappa(-e^{-\kappa L}V_1 + e^{\kappa L}W_1). \tag{3.16}$$

For short expression, k, κ for $A \sim V$ are not shown.

First we erase V_1, W_1, then (A_2, B_2) and be expressed with (A_1, B_1). Because of the linearity, the solution can be written in a matrix form as

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} \equiv M_T \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{3.17}$$

Then matrix $\{m_{ij}\}$ is obtained as

$$\begin{cases} m_{11} = \left[\cosh(\kappa L) + i\dfrac{k^2 - \kappa^2}{2k\kappa}\sinh(\kappa L)\right], \\ m_{12} = -i\dfrac{k^2 + \kappa^2}{2k\kappa}\sinh(\kappa L), \\ m_{21} = m_{12}^*, \qquad m_{22} = m_{11}^*. \end{cases} \tag{3.18}$$

Specific form of M_T surely depends on shape of potential though the relation between input and output can always be written in the matrix form as in (3.17) guaranteed by the linearity of Schrödinger equation. A matrix like M_T is called **transfer matrix** (T-matrix).

In Eq.(3.18), M_T has the symmetry of $m_{21} = m_{12}^*$, $m_{22} = m_{11}^*$, which comes from the time-reversal symmetry and the even symmetry in the potential shape.

Let $B_2 = 0$, and the ratio of transmission wave A_2 and reflection wave B_1 to the incident wave A_1 can be given from (3.17), (3.18) as

$$t \equiv \frac{A_2}{A_1} = \frac{|m_{11}|^2 - |m_{12}|^2}{m_{11}^*} = \frac{1}{m_{11}^*} = \frac{2ik\kappa}{(k^2 - \kappa^2)\sinh(\kappa L) + 2ik\kappa\cosh(\kappa L)}, \tag{3.19}$$

$$r \equiv \frac{B_1}{A_1} = -\frac{m_{21}}{m_{22}} = \frac{(k^2 + \kappa^2)\sinh(\kappa L)}{(k^2 - \kappa^2)\sinh(\kappa L) - 2ik\kappa\cosh(\kappa L)}. \tag{3.20}$$

t, r are called **imaginary transmission coefficient** and **imaginary reflection coefficient** respectively. They are related to the transmission and reflection coefficients as

$$\text{Transmission: } T = |t|^2, \qquad \text{Reflection: } R = |r|^2, \qquad |t|^2 + |r|^2 = 1, \tag{3.21}$$

and the T-matrix M_T can be expressed with them as

$$M_T = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix}. \tag{3.22}$$

3.4.2 Transmission through double-barrier structure

Let us consider the transmission through the double barrier potential illustrated in Fig.3.8. Quantum well and quantum barrier are upside down to each other and the double barrier may have the position in between them. Let the boundaries

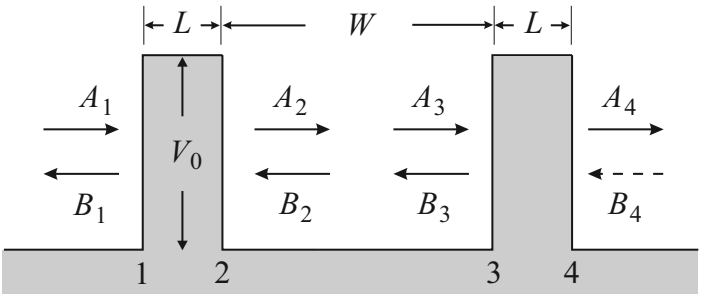


Figure 3.8: Schematic illustration of double barrier potential.

be 1∼4 as in the figure and the wavefunctions also as A_{1-4} and B_{1-4}. For the left barrier the setup is the same as that in the previous section and (3.18) is applcable. Next in the well part between the barriers, a particle gains a kinetic phase factor $\exp(ikW)$ during the traverse. Hence as T-matrix for this part we can adopt

$$M_W = \begin{pmatrix} \exp(ikW) & 0 \\ 0 & \exp(-ikW) \end{pmatrix}. \tag{3.23}$$

The right barrier is just the same as the left. The expression of T-matrix does not depend on local coordinates and M_T can be used as it it.

Then the total T-matrix M_{DW} of the double barrier structure is, as obvious from the definition, obtained as the product of all T-matrices as

$$M_{DW} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} e^{ikW} & 0 \\ 0 & e^{-ikW} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \equiv \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}. \tag{3.24}$$

The transmission coefficient is, from (3.24),

$$T_{11} = m_{11}^2 \exp(ikW) + |m_{12}|^2 \exp(-ikW) \quad (\because m_{12} = m_{21}^*).$$

The interference effect due to the double barrier structure appears in the second term. Let the argument of m_{11} be φ, and writing $m_{11} = |m_{11}| \exp(i\varphi)$ we get

$$\begin{aligned} T_{11}T_{11}^* &= ((|m_{11}|^2 e^{2i\varphi} e^{ikW} + |m_{12}|^2 e^{-ikW})(|m_{11}|^2 e^{-2i\varphi} e^{-ikW} + |m_{12}|^2 e^{ikW}) \\ &= (|m_{11}^2 - |m_{12}|^2)^2 + 2|m_{11}|^2 |m_{12}|^2 \, (1 + \cos(2(\varphi + kW))) \\ &= 1 + 4|m_{11}|^2 |m_{12}|^2 \cos^2(\varphi + kW). \end{aligned}$$

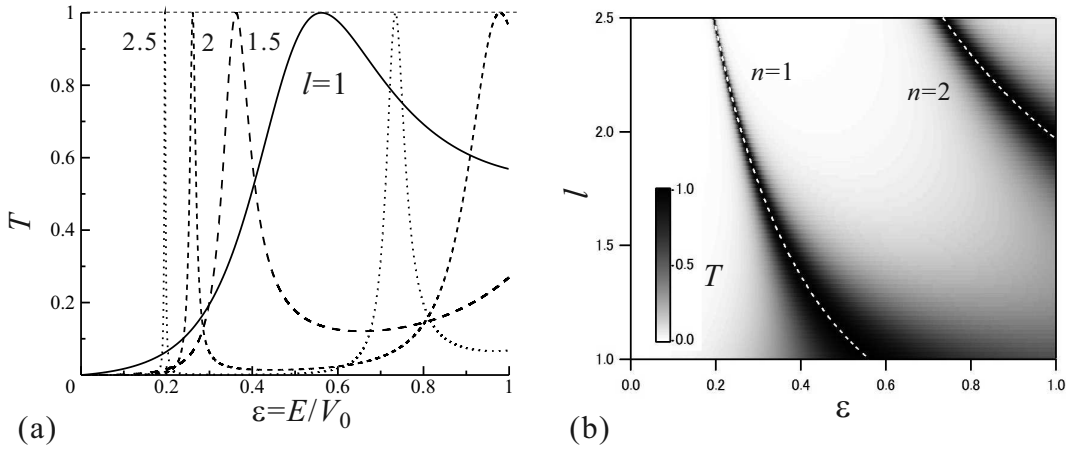


Figure 3.9: (a) Transmission coefficient T calculated on (3.25) as a function of the energy of incoming wave for various barrier widths. Well width - barrier width relation is fixed to $W = 2L$. (b) The same results are plotted in a gray scale as a function of the incoming energy and the barrier width. White broken lines indicate the resonance condition (3.26), (3.27).

The the transmission coefficient is obtained as

$$T = \frac{1}{|T_{11}|^2} = \frac{1}{1 + 4|m_{11}|^2|m_{12}|^2 \cos^2(\varphi + kW)}. \tag{3.25}$$

The final form of transmission coefficient is then in combination obtained with (3.18).

Figure 3.9(a) shows thus calculated transmission coefficient T for various barrier widths L as a function of energy of incoming wave. The relation between the barrier width and the well width is fixed as $W = 2L$. Here L and E are transfomed into dimensionless parameters $l \equiv (\sqrt{2mV_0}/\hbar)L$ and E は $\epsilon \equiv E/V_0$ respectively. The points where the transmission coefficient hits 1 are due to **resonant scattering** and the condition is written as

$$\varphi + kW = \left(n - \frac{1}{2}\right)\pi \qquad (n = 1, 2, \cdots), \tag{3.26}$$

from (3.25), where φ is witten from (3.18) as

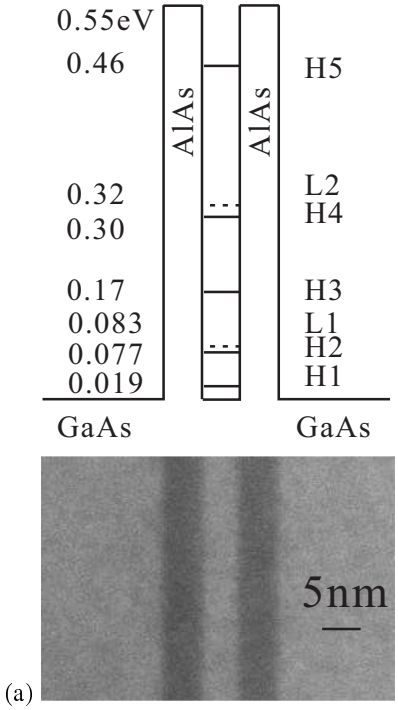
$$\varphi = \arctan\left[\frac{k^2 - \kappa^2}{2k\kappa}\tanh(\kappa L)\right], \tag{3.27}$$

where we restrict the region to $-\pi/2 < \varphi < \pi/2$. With this, n should take a natural number.

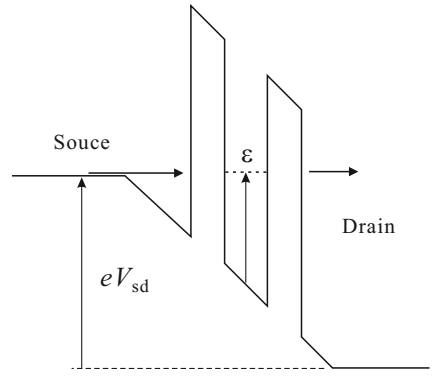
In Fig.3.9(b), the same data are plotted in a gray scale versus a plane of energy and barrier width. White broken lines indicate the resonant scattering condition in the above equation. With increasing l, the peaks become sharper, which tendency is due to the elongation of time for staying inside the well, that makes the life width determined from the uncertainty relation smaller. If we take the limit $L \to \infty$ keeping W finite, the system becomes a quantum well with a finite barrier height and the resonant scattering condition approaches to that for bound eigenstates.

3.4.3 Transport of double barrier diode

Double barrier diode is a device, which realized the double barrier structure with hetero-inferfaces. Here we introduce an experiment on such a device with GaAs-AlAs hetero-interfaces, p-type doped electrodes. Hence the device works as a double barrier for holes. The band discontinuity is $\Delta E_{\rm v}$ =0.47 eV. There are two species of holes at the top of valence band in GaAs with effective masses $0.51m_0$ and $0.082m_0$, which are called "heavy" and "light" holes (hh and lh) respectively. We ignore the mass difference in AlAs for simplicity (actually the difference is not small but does not affect the result significantly). The potential prepared has, as shown in the upper panel of Fig.3.10(a), widths of 5nm



(a)



(b)

Figure 3.10: (a) Upper panel: Potential diagram of the double barrier diode prepared for the present experiment. The energy base is taken to the top of the valence band and the energy of holes is positive in this plot. $H1-H5$, $L1$, $L2$ are the positions of resonant levels for heavy holes and light holes respectively. Lower panel: Scanning transmission electron micrograph of the sample. Darker regions are AlAs. (b) Schematic potential diagram for a source-drain biased double barrier diode.

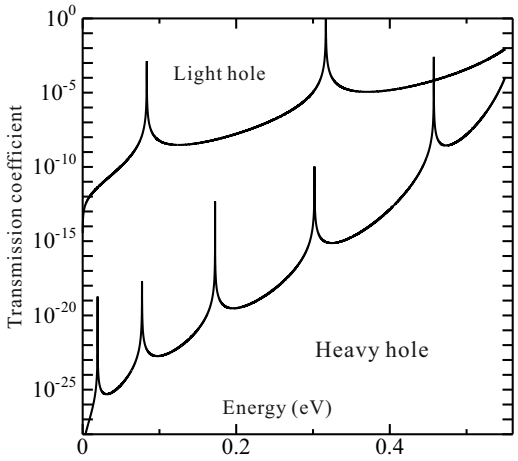


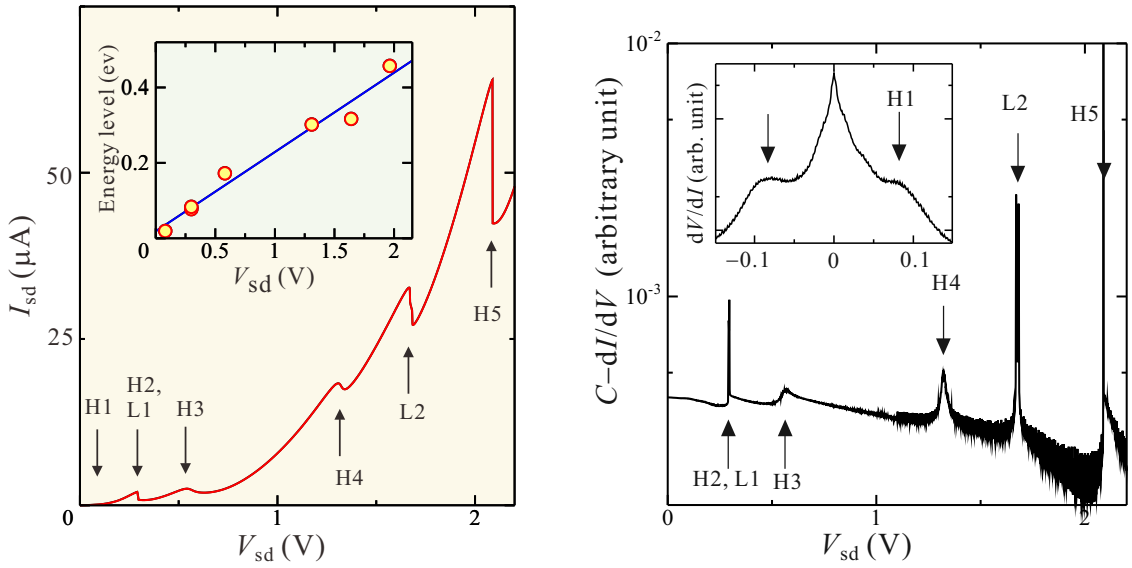
Figure 3.11: Energy dependence of transmission coefficient for the double barrier structure with parameters given in the text and with eq.(3.25). The peaks hit 1 actually but too narrow to be sampled.

both for the barriers and the well. The barriers and the well parts do not have any doping. Figure 3.10(a) shows a photograph of the sample cross section taken by a scanning transmission electron microscope, STEM.

The transmission coefficient T thus calculated with the above parameters and the structure shown in Fig.3.10 is displayed as a function of energy in Fig.3.11. Because the effective masses of holes are comparatively heavy and the barrier height is high, the transmission peaks are very sharp. We thus can see the behavior of tail only in the semi-log plot. We see below the barrier threshold, 5 heavy hole resonance peaks and 2 light hole peaks. Figure 3.10(a) shows the positions of resonance levels in the well numerically calculated from eq.(3.26).

In order to see the behavior of tunneling, usually source-drain voltage V_{sd} is applied as illustrated in Fig.3.10(b). Inside the source and the drain, highly concentrated holes screen the electric field and the applied voltage should mainly consumed across the double barrier regions. In actual situation, however, the contact resistances also cause significant voltage drops.

We ignore the distortion of the originally rectangular-shaped potential due to the applied electric field. Then, as in the illustration, the energy of an injected hole is in accordance with the resonant level when the applied voltage reaches twice of it. The transmission coefficient takes a peak at that time, that is the amount of holes passing through the barriers within a unit time, namely the current should take a peak (see Appendix E for more realistic current lineshape).



(a) (b)

Figure 3.12: (a) Current-voltage characteristics of the double barrier diode introduced in Fig.3.10. Resonant levels corresponding to the peaks are indicated by the arrows. The inset indicates peak positions of energy levels on the voltage axis. (b) Emphasis is on the peak positions with differentiating the current with the voltage and the absolute value being plotted in semi-log scale. The inset is enlargement around the origin.

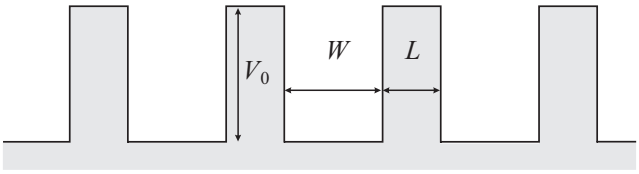


Figure 3.13: One dimensional rectangular potential (Kronig-Penny type potential)

A measured current-voltage curve in a double barrier diode (the one in Fig.3.10) is shown in Fig.3.12(a). Several current peaks appear versus the voltage. To clarify the peak positions the absolute value of voltage-derivative the current with a constant bias C is plotted in a semi-log scale in Fig.3.12(b) [2].

3.4.4 Superlattice

The next step, in the course of quantum mechanics, we have double quantum well, which is very important as a qubit. We skip it, to my regret, for the shortage of time. I would like to remind you we have lectures on "nano-quantum information" in the applied physics department (but in Japanese). Here I would like to give a short introduction of **heterojunction superlattice**, which was proposed by Leo Esaki and Raphael Tsu and has provided rich physics. The basic idea of heterojunction superlattice is realization of Kronig-Penny type potential, illustrated in Fig.3.13. This, in a sense, recovers spatial translational symmetry of the lattice lost by the introduction of the interface but in a different manner.

Let us express a Kronig-Penny type potential as $V_{\rm KP}(x)$ and write down the Schrödinger equation as

$$\left[-\frac{\hbar^2 d^2}{2mdx^2} + V_{\rm KP}(x)\right]\psi(x) = E\psi(x), \qquad V_{\rm KP}(x) = V_{\rm KP}(x+d). \tag{3.28}$$

According to Bloch theorem, we write the eigenstate wavefunction as a product of a plane wave and a lattice periodic function with $d = L + W$ as the lattice constant.

$$\psi_K(x) = u_K(x)e^{iKx}, \quad u_K(x+d) = u_K(x), \quad K \equiv \frac{\pi s}{Nd}. \tag{3.29}$$

s takes an integer from $-N+1$ to $N-1$. The transfer matrix M_d corresponding to the unit cell of the system is

$$M_d = \begin{pmatrix} e^{ikW} & 0 \\ 0 & e^{-ikW} \end{pmatrix}\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} m_{11}e^{ikW} & m_{12}e^{ikW} \\ m_{21}e^{-ikW} & m_{22}e^{-ikW} \end{pmatrix}. \tag{3.30}$$

As before, we write the input/output in the left hand side of i-th cell as (a_i, b_i), then from (3.29),

$$\begin{pmatrix} a_{i+1} \\ b_{i+1} \end{pmatrix} = M_d \begin{pmatrix} a_i \\ b_i \end{pmatrix} = e^{iKd}\begin{pmatrix} a_i \\ b_i \end{pmatrix} \tag{3.31}$$

should hold, that is, this is a problem of engenvalue e^{iKd} of matrix M_d. From the unitarity of M_d, or from "reversed" equation of (3.31), the two eigenvalues $e^{\pm iKd}$ are obtained. We re-use $\{m_{ij}\}$ in (3.18) to get to the equations

$$e^{iKd} + e^{-iKd} = 2\cos Kd = {\rm Tr}M_d = 2{\rm Re}(e^{-ikW}m_{11}^*), \tag{3.32}$$

$$\cos\left[K(L+W)\right] = \cosh(\kappa L)\cos(kW) - \frac{k^2-\kappa^2}{2k\kappa}\sinh(\kappa L)\sin(kW). \tag{3.33}$$

By use of φ in (3.27), expression

$$\cos(Kd) = |m_{11}|\cos(kW+\varphi) = \frac{1}{|t|}\cos(kW+\varphi) \tag{3.34}$$

is available.

Transforming the Kronig-Penny potential to a series of δ-function potentials can be attained with taking limits $L \to 0$, $W \to d, V_0 \to \infty (V_0 L = C$(constant) to obtain the condition

$$\cos(Kd) = \cos(kd) + \frac{mC}{\hbar^2 k}\sin(kd). \tag{3.35}$$

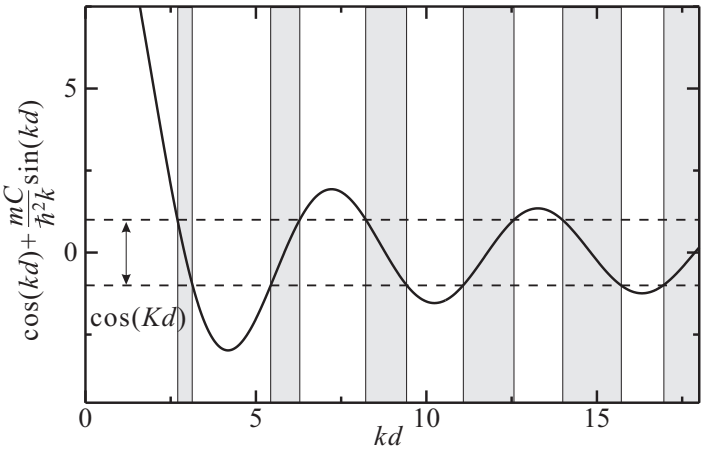


Figure 3.14: RHS of (3.35) as a function of kd. Here $mdC/\hbar^2$ is taken to be 13. The gray belts indicate "allowed bands".

Figure 3.14 shows the RHS as a function of kd. The solution K for (3.35) exists for the RHS to be in $[-1, +1]$ corresponding to the gray bands namely the energy bands.

Let us simplify the energy dispersion relation of a single band as

$$E(K) = \frac{E_{nw}}{2}(1 - \cos Kd). \tag{3.36}$$

The group velocity and the effective mass are

$$v_g(K) = \frac{E_{nw}d}{2\hbar}\sin Kd, \qquad m^*(K) = \frac{\hbar^2}{E_{nw}d^2}\sec Kd. \tag{3.37}$$

The equation of motion of an electron in a periodic potential under a uniform electric field E_m is written as

$$m^*\frac{dv}{dt} = \hbar\frac{dK}{dt} = F = eE_m. \tag{3.38}$$

We see an effective mass in a periodic potential can be negative.

An acceleration according to (3.38) results in $K = eE_m t/\hbar$. Now we put a wave packet with zero-velociy at the origin $x = 0$, and observe the time evolution. From (3.37),

$$v_g(t) = \frac{E_{nw}d}{2\hbar}\sin\left(\frac{eE_m d}{\hbar}t\right), \qquad x(t) = \frac{E_{nw}}{2eE_m}\left[1 - \cos\left(\frac{eE_m d}{\hbar}t\right)\right]. \tag{3.39}$$

The result indicates that in spite of the constant acceleration, the wave packet oscillates in space. The phenomenon is called **Bloch oscillation**, an observation of which in an actual lattice is almost impossible due to various scattering. In a superlattice, however, the super-period devides the large original band into "mini-bands" and the acceleration to the top of a mini-band before scattering. The Block oscillation was thus observed in superlattices in optical measurements.

3.5 Modulation doping and two-dimensional electrons

The most popuar artificial structure made with heterojunctions is the two-dimensional electrons with modulation doped heterojunctions (two-dimensional electron gas, 2DEG). As is illustrated in Fig.3.15, in a single heterojunction, doping is given just in the wider band region. Now let us see what happens here for n-type doping.

Let us take the z-axis vertical to the surface and the hetero-interface plane as in the figure. In a "rigid band" model, the conduction band discontinuity ΔE_c emerges and the carriers re-distribute. Let us take the plain case of the combination of $Al_xGa_{1-x}As$ and GaAs. Then we can adopt the approximation that the envelope function in the effective mass approximation as the electron wavefunction itself, and electron-electron interaction can be treated within the Hartree approximation[3]. Then, the Poission-Schrödinger equation including the electrostatic potential formed by ionized donor, the band discontinuity and the 2DEG itself should be solved self-consistently for obtaining quantilzed energy levels and wavefunction (envelope function) along the direction perpendicular to the 2DEG plane.

[2]This transformation is just for the clarity in sight.

[3]Even within the mean field theory, the interaction term contains the Fock term (exchange), but the contribution was calculated to be small.

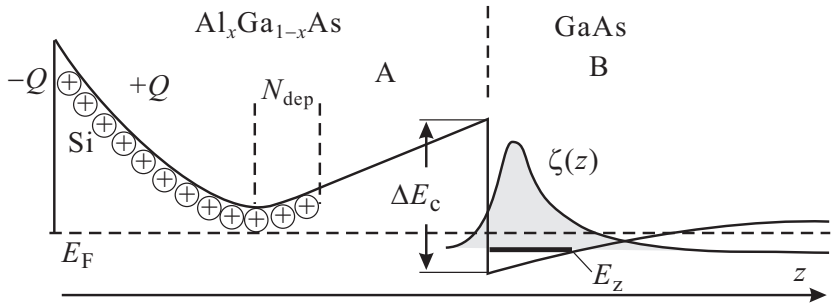


Figure 3.15: Schematic cross sectional view of two-dimensional electrons at a modulation doped $Al_xGa_{1-x}As$/GaAs heterointerface.

z-axis is taken to be perpendicular to the heterointerface plane. As in Fig.3.15, the surface Schottky barrier creates a surface depletion layer. Let the charge at the surface be $-Q$ and the electric field from the charge should be compensated with that from charges at ionized donors (in the figure Si) $+Q$ in the amount and screened from inside. Let us write the number of all the residual ionized donors per unit area (integrated along z-axis) as $N_{\rm ddep}$. The electrostatic potential from the charges is, far inside the lattice from the doping region, $V_D(z) = (4\pi e^2/\epsilon\epsilon_0)N_{\rm dep}z$. Between the doped region and the hetero-interface, a non-doped region called "spacer " is often places. The spacer spatially separates the 2DEG and the ionized impurities, decreases scattering probabilities of two-dimensional electrons, resulting in very high mobility of electrons. A too thick spacer, however, lifts up the band depletes the well and throws out the 2DEG.

Let us adopt a variable separation type expression for 2DEG wavefuntion, $\Psi(\boldsymbol{r}) = \psi(x,y)\zeta(z)$. $\zeta(z)$ is the envelope function along z-axis. The areal concentration n_{2d} is the function of discretized energy level E_z, which is in other words the kinetic energy along z-axis for $\zeta(z)$. The areal charge density at position z' is then $-en_{2d}|\zeta(z')|^2$, the sheet charge of which creates the electric field $-(4\pi e^2/\epsilon\epsilon_0)n_{2d}|\zeta(z')|^2|z-z'|$ as calculated from the Gauss theorem. In the Hartree-only mean field approximation, the potential should include these terms. The potential created by the 2DEG itself is

$$V_{2d}(z) = -\frac{4\pi e^2}{\epsilon\epsilon_0}n_{2d}(E_z)\int_{-\xi}^{\infty}|\zeta(z')|^2|z-z'|dz'.$$

Here the integral cut-off ξ should be taken longer enough than the penetration depth of $\zeta(z)$ in to AlGaAs barrier. We write a step potential with discontinuity ΔE_c just at the interface as $V_h(z)$. Now the total potential can be written as

$$V(z) = V_h(z) + \frac{4\pi e^2}{\epsilon\epsilon_0}\left[N_{\rm dep}z - n_{2d}(E_z)\int_{-\xi}^{\infty}|z-z'||\zeta(z')|^2dz'\right]. \tag{3.40}$$

Schrödinger equation for $\zeta(z)$

$$\left[-\frac{\hbar^2}{2m^*(z)}\frac{\partial^2}{\partial z^2} + V(z)\right]\zeta(z) = E_z\zeta(z) \tag{3.41}$$

should be solved self-consistently to obtain (consistent) $\zeta(z)$. The effective masses m^* are different in the two species of semiconductors and the boundary condition should be

$$\zeta(0)^{(\rm A)} = \zeta(0)^{(\rm B)}, \quad \left.\frac{1}{m^*_{\rm A}}\frac{d\zeta^{(\rm A)}}{dz}\right| = \left.\frac{1}{m^*_{\rm B}}\frac{d\zeta^{(\rm B)}}{dz}\right|. \tag{3.42}$$

In the Poisson-Schrödinger procedure, one should solve the equations from (3.40) to (3.42) consistently. The above only treats the Hartree term. In general, the Fock term, or the correlation effect is also important in mean field theory. However, it is known that the correlation effect does not affect $\zeta(z)$ or E_z so much and here we ignore it for simplicity.

It is comparatively easy to solve Poisson-Schrödinger equation numerically for a simple band with small spin-orbit interaction, like the conduction band in GaAs. For more complicated cases, *e.g.*, multiple valleys, strong spin-orbit interaction, etc., the scale of numerical calculation increases. If one needs to expand the calculation to other quantities with obtained $\zeta(z)$ for such a case, approximate formulas with simple mathematical forms are convenient. For example, in Fang-Howard approximation, the formula

$$\zeta(z) = \sqrt{\frac{b^2}{2}}z\exp\left(-\frac{bz}{2}\right) \tag{3.43}$$

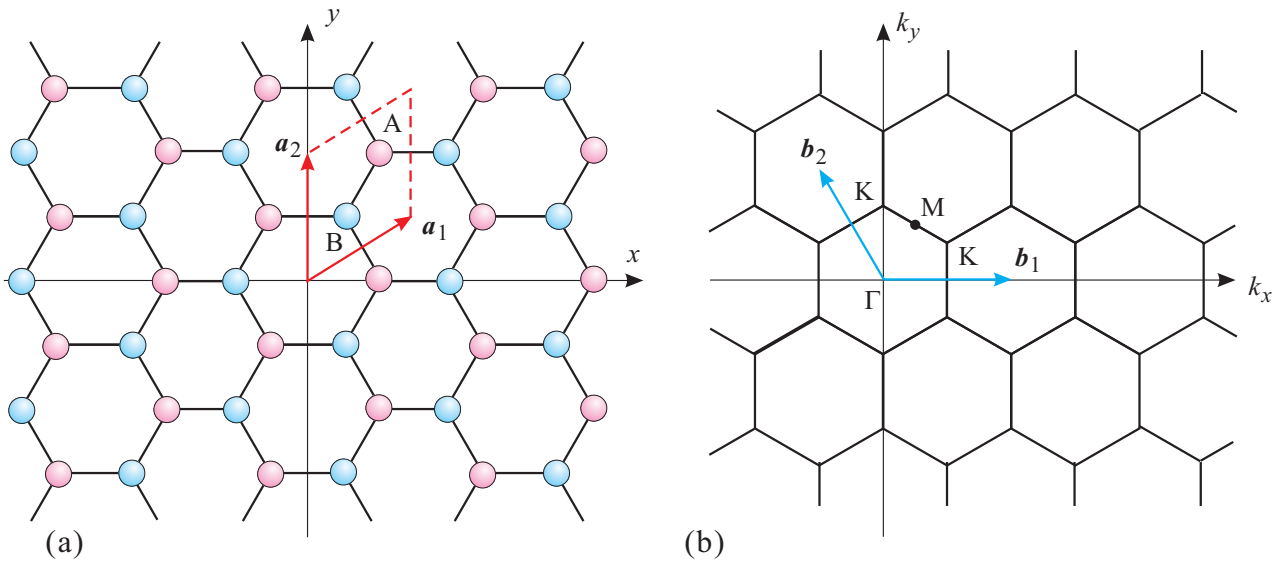


Figure 3.16: (a) Two-dimensional crystal structure in graphene. Carbon atoms form a two-dimensional honeycomb lattice. This can be also viewed as a superposition of two face-centered rectangular lattices (A position atoms and B position atoms). (b) Reciprocal lattice of the lattice in (a). $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are reciprocal lattice vectors corresponding to $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ respectively. Γ-point is the center of first Brillouin zone and symmetric points K-point and M-point are also indicated in the figure.

is used as the trial function with b as a parameter for variational calculation. The result of the variational calculation is given as

$$b^3 = \frac{48\pi m e^2}{\epsilon\epsilon_0\hbar^2}\left(\frac{11}{32}n_{2d} + N_d\right). \tag{3.44}$$

In this approximation, penetration of wavefunction into the barrier (spacer) is ignored. Another approximation form which takes such penetration into account is given in, *e.g.* ref.[6].

3.6 A two-dimensional material: band structure of graphene

Another way to form two-dimensional electron system, is the utilization of two-dimensional materials, in which the atoms aligned on two-dimensional planes. Graphene is a representative two-dimensional material, in which in-plane atomic connection is on sp^2 σ-bonding and a kind of three-way standoff appears in the structure resulting in so-called Dirac point.

As shown in Fig.3.16(a), single layer graphene crystal is a two-dimensional honeycomb lattice of carbon. The diamond shown in the figure is a unit cell and unit lattice vectors and reciprocal vectors are

$$\boldsymbol{a}_1 = \begin{pmatrix}\sqrt{3}a/2\\ a/2\end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix}0\\ a\end{pmatrix}, \quad \boldsymbol{b}_1 = \begin{pmatrix}4\pi/\sqrt{3}a\\ 0\end{pmatrix}, \quad \boldsymbol{b}_2 = \begin{pmatrix}-2\pi/\sqrt{3}a\\ 2\pi/a\end{pmatrix}. \tag{3.45}$$

Below we calculate electronic states of graphene under simplest tight binding model. It is a coarse approximation, which cannot be compared with experiments quantitatively. It may be a help, however, to understand why Dirac points are placed at the Fermi energy in graphene. Carbon is a group IV element and the outermost electrons are in $2s$, $2p_x$, $2p_y$, $2p_z$. From the view of chemical bonds, as can be guessed from the lattice structure, taking linear combinations of these orbitals, they are divided into σ electrons in sp^2 hybrid orbitals and π electrons. σ electrons are in covalent bonds and the energy band places at a low position forming the honeycomb lattice. Hence the Fermi level is determined by π electrons. Let us consider, then, the π-electrons on a honeycomb lattice and write down Schrödinger equation.

From

$$\psi = \mathscr{H}\psi, \tag{3.46}$$

we divide the wavefunction to those on site A and B as shown in Fig.3.16(a) and apply tight-binding approximation.

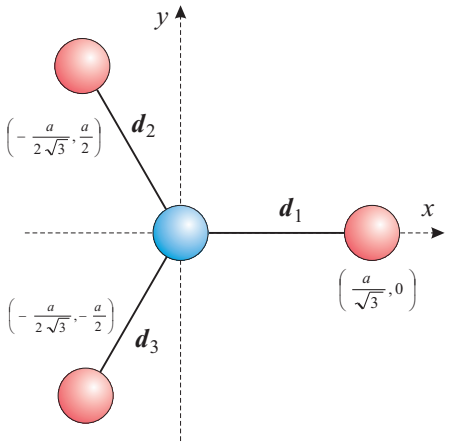


Figure 3.17: Three vectors representing the nearest neighbour coupling of a carbon atom.

That is

$$\psi = \zeta_{\mathrm{A}}\psi_{\mathrm{A}} + \zeta_{\mathrm{B}}\psi_{\mathrm{B}}, \tag{3.47}$$

$$\psi_{\mathrm{A}} = \sum_{j\in A} \exp(i\boldsymbol{k}\boldsymbol{r}_j)\phi(\boldsymbol{r}-\boldsymbol{r}_j), \tag{3.48a}$$

$$\psi_{\mathrm{B}} = \sum_{j\in B} \exp(i\boldsymbol{k}\boldsymbol{r}_j)\phi(\boldsymbol{r}-\boldsymbol{r}_j), \tag{3.48b}$$

where $\phi(\boldsymbol{r})$ is an atomic wavefunction of π-electron, $\boldsymbol{r}_j$ is a lattice point. We prepare wavefunctions on partial lattices, Hamiltonians for partial lattices and between the lattices as

$$H_{\mathrm{AA}} = \langle\psi_{\mathrm{A}}|\mathscr{H}|\psi_{\mathrm{A}}\rangle, \quad H_{\mathrm{BB}} = \langle\psi_{\mathrm{B}}|\mathscr{H}|\psi_{\mathrm{B}}\rangle, \quad H_{\mathrm{AB}} = H^*_{\mathrm{BA}} = \langle\psi_{\mathrm{A}}|\mathscr{H}|\psi_{\mathrm{B}}\rangle. \tag{3.49}$$

The total number of atoms in the system is taken as $2N$, that is

$$\langle\psi_{\mathrm{A}}|\psi_{\mathrm{A}}\rangle = \langle\psi_{\mathrm{B}}|\psi_{\mathrm{B}}\rangle = N. \tag{3.50}$$

In the tight binding model ($\langle\psi_{\mathrm{A}}|\psi_{\mathrm{B}}\rangle = 0$), eq.(3.47) is substituted into (3.46) and the condition for the existence of non-trivial solution of $(\zeta_{\mathrm{A}}, \zeta_{\mathrm{B}})$ leads to the secular equation

$$\begin{vmatrix} H_{\mathrm{AA}} - NE & H_{\mathrm{AB}} \\ H_{\mathrm{BA}} & H_{\mathrm{BB}} - NE \end{vmatrix} = 0. \tag{3.51}$$

Then we obtain

$$E = (2N)^{-1}\left(H_{\mathrm{AA}} + H_{\mathrm{BB}} \pm \sqrt{(H_{\mathrm{AA}} - H_{\mathrm{BB}})^2 + 4|H_{\mathrm{AB}}|^2}\right) \equiv h_{\mathrm{AA}} \pm |h_{\mathrm{AB}}|, \tag{3.52}$$

where we have used the symmetry $H_{\mathrm{AA}} = H_{\mathrm{BB}}$, and the small letter quantities are atomic ones obtained by multiplying $(2N)^{-1}$.

Now for

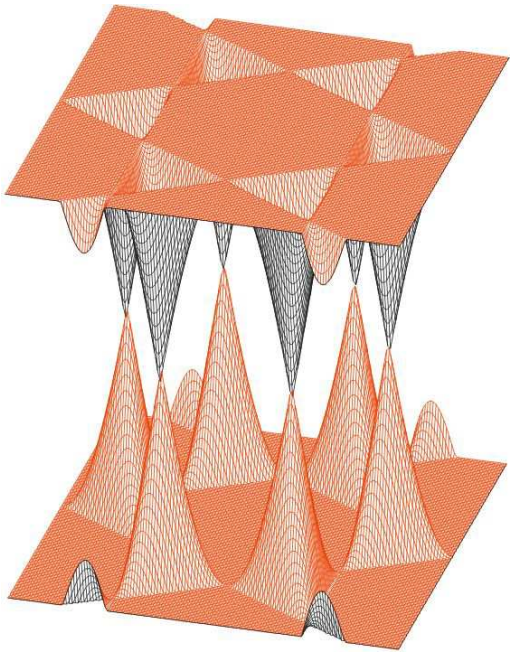
$$H_{\mathrm{AB}} = \sum_{l\in\mathrm{A}, j\in\mathrm{B}} \exp\left[i\boldsymbol{k}(\boldsymbol{r}_j - \boldsymbol{r}_l)\right] \langle\phi(\boldsymbol{r}-\boldsymbol{r}_l)|\mathscr{H}|\phi(\boldsymbol{r}-\boldsymbol{r}_j)\rangle_{\boldsymbol{r}}, \tag{3.53}$$

we apply further approximation that the integral $\langle\phi(\boldsymbol{r}-\boldsymbol{r}_l)|\mathscr{H}|\phi(\boldsymbol{r}-\boldsymbol{r}_j)\rangle_{\boldsymbol{r}}$ is non-zero only for the nearest neighbour. For calculation of a single term, we pick up the atom labeled A in Fig.3.16(a), and the vectors to the nearest neighbour 1, 2, 3 as$\boldsymbol{d}_i (i = 1, 2, 3)$. As we can see in the figure,

$$\boldsymbol{k}\cdot\boldsymbol{d}_1 = \frac{k_x a}{\sqrt{3}}, \quad \boldsymbol{k}\cdot\boldsymbol{d}_2 = \left(-\frac{k_x}{2\sqrt{3}} + \frac{k_y}{2}\right)a, \quad \boldsymbol{k}\cdot\boldsymbol{d}_3 = \left(-\frac{k_x}{2\sqrt{3}} - \frac{k_y}{2}\right)a, \tag{3.54}$$

where $a = |\boldsymbol{a}_1| = |\boldsymbol{a}_2|$. And the terms $\langle\phi(\boldsymbol{r}-\boldsymbol{r}_l)|\mathscr{H}|\phi(\boldsymbol{r}-\boldsymbol{r}_j)\rangle_{\boldsymbol{r}}$ should be the same for the symmetry and we write the value as ξ. Then we obtain

$$h_{\mathrm{AB}} = \left(\sum_{j=1}^{3} \exp(i\boldsymbol{k}\cdot\boldsymbol{d}_j)\right)\xi. \tag{3.55}$$



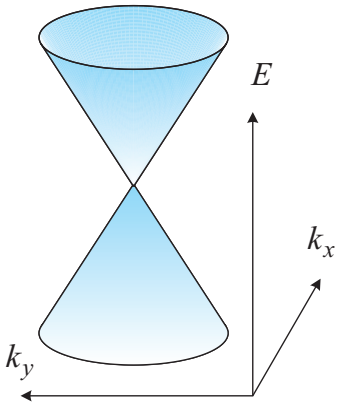


Figure 3.18: Left panel: Three dimensional wireframe plot of energy band in graphene (3.56). The dispersion around the Dirac points are displayed. Upper and lower energy parts are cut off for convenience. Upper panel : Illustration of Dirac corn.

We obtain the next expression for the eigenenergy with substituting eqs.(3.54), (3.55) into (3.52).

$$E = h_{\rm AA} \pm \xi \sqrt{1 + 4\cos\frac{\sqrt{3}k_x a}{2}\cos\frac{k_y a}{2} + 4\cos^2\frac{k_y a}{2}}. \tag{3.56}$$

The second term is the perturbation from resonant integral between neighboring atoms. It vanishes at K-point in the reciprocal lattice, *i.e.*,

$$(k_x, k_y) = \left(0, \pm\frac{4\pi}{3a}\right), \quad \left(\frac{2\pi}{\sqrt{3}a}, \pm\frac{2\pi}{3a}\right), \quad \left(-\frac{2\pi}{\sqrt{3}a}, \pm\frac{2\pi}{3a}\right). \tag{3.57}$$

Fixing k_y to $k_y = 4\pi/3a$, and around $k_x = 0$(a K-point), (3.56) gives

$$E\left(k_x, \frac{4\pi}{3a}\right) \approx h_{\rm AA} + \frac{\sqrt{3}\xi a}{2}|k_x|. \tag{3.58}$$

Namely, the upper band at a K-point shows linear dispersions which form a downward cusp at zero energy. In the lower again linear dispersions form an upward cusp, hence resulting in so called **Dirac point** at the K-point, which does not have energy gap nor effective mass.

The details in (3.56) cannot be compared with a realistic band structure, but just like cosine band in one-dimensional tight-binding model, we can easily have insight into, *e.g.*, why Dirac points appear at K-points. It is apparent that the three fold symmetry of the resonant integral in (3.55) is the origin of the Dirac points. The symmetry does not change with the accuracy of the approximation and we can conclude that the Dirac points exist at K-points in real graphenes. It is also an interesting example that such a symmetry causes degeneracy in energy levels.

References

[1] R. L. Anderson, IBM J. Res. Dev. **4**, 283 (1960).

[2] J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).

[3] H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers Part B: Materials and Operating Characteristics* (Academic Press, 1978).

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

[5] L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 61 (1970).

[6] T. Ando, J. Phys. Soc. Jpn. **51**, 3893 (1982).

Appendix D: Excitons in low dimensional systems

An exciton is a bound state of a hole and an electron created by photon irradiation. Excitons have great effects on optical absorption and emission. An exciton is more stable than a free electron-hole pair by the binding energy, has a longer lifetime, hence the absorption coefficient is larger than that of the band-edge. There are many types of excitons and here we consider Wannier-type free excitons, which have much larger spatial size than the lattice constant. Wannier-type free exciton can be treated as a single-body problem within a simple approximation. Let the effective binding potential be $V(\boldsymbol{r}) = a/|\boldsymbol{r} - \boldsymbol{r}_0|$. Bounded energy levels E_n $(n = 1, 2, \cdots)$ are hydrogen-atom like with the effective Rydberg constant R^*. Namely they are written as $E_n = -R^*/n^2$. In quantum confined structures, when the confinement length is shorter than the effective Bohr radius $a_{\rm B}^*$, the dimension for excitons are also lowered.

Let us treat it as a problem of a hydrogen atom then we treat Schrödinger equation with a Coulomb-type central force potential $V_{\rm c}(\boldsymbol{r})$,

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 + V_{\rm c}(\boldsymbol{r})\right)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{D.1}$$

in lower dimensions. Here m^* is the electron-hole reduced mass. And we need to change the potential form as

$$V_{\rm c}^{\rm 2d}(\boldsymbol{r}) = -\frac{e^2}{4\pi\epsilon\epsilon_0|\boldsymbol{r}|}, \quad V_{\rm c}^{\rm 1d}(r) = -\frac{e^2}{4\pi\epsilon\epsilon_0(|z| + 0.3r_0)}, \tag{D.2}$$

particularly for one-dimensional (along z-axis) systems. This is because simple transformation of eq.(D.1) into one-dimension causes anomalous behavior including divergence of binding energy. The potential form in eq.(D.2) is given as an empirical formula which well fits to a practical numerical calculation on confinement into a finite width quantum wire (a cylinder with radius r_0). Below, we rapidly see the solutions, which are nothing but hydrogen atom solutions. Under variable separation hypothesis, the solutions for eq.(D.1) can be written in the forms

$$\psi^{\rm 3d} = \rho^l e^{-\rho/2} R(\rho) Y_{l,m}(\theta, \varphi), \quad \psi^{\rm 2d} = \rho^{|m|} e^{-\rho/2} R(\rho) e^{im\varphi}, \quad \psi^{\rm 1d} = R(\zeta). \tag{D.3}$$

ρ and ζ are dimensionless variables, which correspond to radial variable and z variable respectively. The definitions are

$$\rho = \alpha r, \quad \zeta = \alpha(|z| + 0.3r_0), \quad \alpha = \frac{\sqrt{-8m^*E}}{\hbar}. \tag{D.4}$$

$R(\rho)$, $R(\zeta)$ are the solutions of the following equations.

$$\begin{cases} \left(\rho\dfrac{\partial^2}{\partial\rho^2} + (p+1-\rho)\dfrac{\partial}{\partial\rho} + q\right)R(\rho) = 0: & \text{3-, 2-dimensional,} \\ \left(\dfrac{\partial^2}{\partial\zeta^2} + \dfrac{\partial}{\partial\zeta} + \dfrac{\lambda}{\zeta}\right)R(\zeta) = 0, \ \lambda \equiv \dfrac{e^2}{4\pi\epsilon_0\hbar}\sqrt{-\dfrac{m^*}{2E}}: & \text{1-dimensional,} \end{cases} \tag{D.5}$$

where p, q are

$$p = \begin{cases} 2l+1 & \text{(3-dimensional)} \\ 2|m| & \text{(2-dimensional)} \end{cases}, \quad q = \begin{cases} \lambda - l - 1 & \text{(3-dimensional)} \\ \lambda - |m| - 1/2 & \text{(2-dimensional)} \end{cases}, \tag{D.6}$$

where l is angular momentum quantum number and m is magnetic quantum number.

For three and two dimensional systems, $R(\rho)$ in eq.(D.5) is expanded as follows.

$$R(\rho) = \sum_\nu \beta_\nu \rho^\nu, \quad \beta_{\nu+1} = \beta_\nu \frac{\nu - q}{(\nu+1)(\nu+p+1)}. \tag{D.7}$$

For this $R(\rho)$ to be finite, this expansion should stop at a finite number, which condition requires $\nu_{\max} = q$. The main quantum number q then is defined as follows.

$$n \equiv \lambda = \nu_{\max} + l + 1 \ \text{ (3-dimensional)}, \quad n \equiv \lambda - \frac{1}{2} = \nu_{\max} + |m| \ \text{ (2-dimensional)}. \tag{D.8}$$

The exciton energy levels for three- and two-dimensional systems can be expressed as follows.

$$E_{{\rm b}n}^{\rm 3d} = -\frac{E_0}{n^2} \qquad n = 1, 2, \cdots, \tag{D.9}$$

$$E_{{\rm b}n}^{\rm 2d} = -\frac{E_0}{(n+1/2)^2} \qquad n = 0, 1, \cdots. \tag{D.10}$$

Here the energy unit E_0 is

$$E_0 = \frac{e^2}{8\pi\epsilon\epsilon_0 a_0^*}, \quad a_0^* = \frac{4\pi\epsilon\epsilon_0\hbar^2}{m^* e^2}, \tag{D.11}$$

where a_0^* is the effective Bohr radius. From eq.(D.8), we see that $n = 0$ is available for two-dimensional systems and the ground bound state energy is $-4E_0$. This means the binding energy is four times larger than that in three-dimensional systems where the ground state energy is $-E_0$. In the process of an exciton formation, spatial confinement increases the kinetic energy due to the uncertainty in momentum. In three-dimensional systems, the enhancement occurs for all three dimensions while in two dimensional systems, the confinement along the direction perpendicular to the plane has already been included into the shift of band edge and the binding energy is measured from the edge. Hence it is qualitatively easily understood that the exciton binding energy becomes larger with lowering the system dimension.

Generally radial wavefunction is expressed with Laguerre bi-polynomial and exponential functions. In three dimensional systems, $1s$ wavefunction is written as $\psi_{1s}^{3d} \propto \exp(-r/a_0^*)$. Similarly let $\psi_{1s}^{2d} \propto \exp(-r/a_0^{*2d})$, (D.5) $\curvearrowleft l = m = 0$ and substitution into Schrödinger equation gives $a_0^{*2d} = a_0^*/2$. The spatial size of excitons in two-dimensional systems is half of that in three-dimensional systems in accordance with increment in the binding energy.

Appendix E: Current-voltage characteristics in double barrier resonant tunneling device

Acutual electric transport in a resonant diode is affected by various factors but in reality, it is recommended that calculation of the transmission coefficiency in a simple model and taking other effects into account in the next step. Let us first decompose the enery of incoming electrons into component perpendicular to the barrier E_z and that parallel to the barrier $E_\parallel$. We ignore anisotropy of the effective mass. The current flows from the left electrode to the right one can be written with the group velocity $v_{gz} = \partial E/\hbar\partial k_z$ along z-axis as

$$\begin{aligned} J_{\mathrm{L}\to\mathrm{R}} &= e\sum_k v_{gz} f_{\mathrm{L}}(1-f_{\mathrm{R}})\mathcal{T} \\ &= \frac{2e}{(2\pi)^3\hbar}\int\int \mathrm{d}^{(2)}k_\parallel \mathrm{d}k_z \left(\frac{\partial E_z}{\partial k_z}\right) f_{\mathrm{L}}(1-f_{\mathrm{R}})\mathcal{T} \\ &= \frac{em}{2\pi^2\hbar^3}\int_0^\infty\int_0^\infty \mathrm{d}E_z \mathrm{d}E_\parallel f_{\mathrm{L}}(E)(1-f_{\mathrm{R}}(E))\mathcal{T}(E_z). \end{aligned} \tag{E.1}$$

Here we assume the dispersion of electrons can be described with a simple parabollic band with no anomaly. $\mathcal{T}(E_z)$ is the transmission coefficient of the barrier at energy E_z (the notation is a bit changed to avoid confusion with temperature.)

The total current flow J is obtained subtracting the counter flow from the above $J_{\mathrm{L}\to\mathrm{R}}$ and given as

$$J = (J_{\mathrm{L}\to\mathrm{R}} - J_{\mathrm{R}\to\mathrm{L}}) = \int_0^\infty \mathrm{d}E_z \mathcal{T}(E_z) S(E_z) \tag{E.2}$$

$$S(E_z) = \frac{em}{2\pi^2\hbar^3}\int_0^\infty \{f_{\mathrm{L}}(E) - f_{\mathrm{R}}(E)\}\, dE_\parallel. \tag{E.3}$$

$S(E_z)$ is called supply function. We adopt a Fermi-Dirac distribution for f and obtain S with writing $\beta = (k_{\mathrm{B}}T)^{-1}$ as

$$S = \left(\frac{emk_{\mathrm{B}}T}{2\pi^2\hbar^3}\right)\ln\left[\frac{1+\exp\beta(E_{\mathrm{F}}-E_z)}{1+\exp\beta(E_{\mathrm{F}}-E_z-eV)}\right]. \tag{E.4}$$

When the system is Fermi-degenerated

$$S(E_z) = \begin{cases} (em/2\pi^2\hbar^3)(E_{\mathrm{F}}-E_z) & (E_{\mathrm{F}}-V \le E_z \le E_{\mathrm{F}}) \\ (em/2\pi^2\hbar^3)eV & (0 \le E_z \le E_{\mathrm{F}}-eV) \end{cases}, \tag{E.5}$$

that is a trapezoidal function. From the above we can calculate J for the case $eV < E_{\mathrm{F}}$ as

$$J = \frac{em}{2\pi^2\hbar^3}\left[eV\int_0^{E_{\mathrm{F}}-eV} dE_z T(E_z) + \int_{E_{\mathrm{F}}-eV}^{E_{\mathrm{F}}} dE_z (E_{\mathrm{F}}-E_z)T(E_z)\right]. \tag{E.6}$$

Though for detailed comparison with experiments, there are many other factors to be accounted, here we shift to a rough approximation. S has a trapezoidal form as shown in Fig.3.19, and the form is transformed into a triangle when

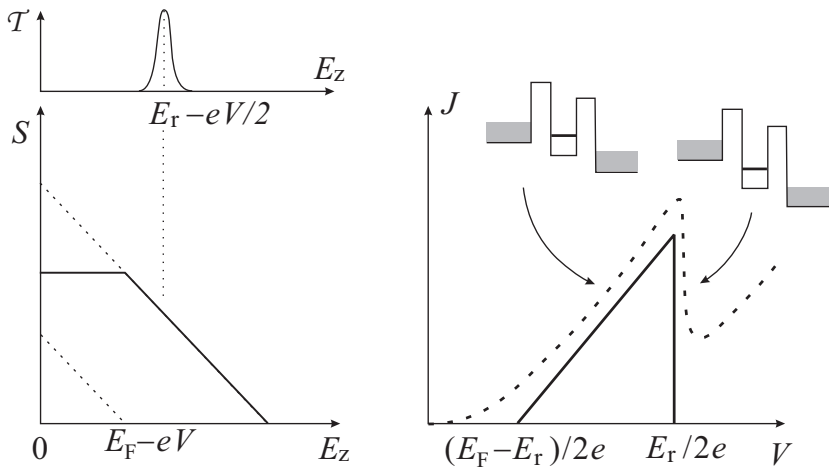


Figure 3.19: Left panel: Schematic illustration of a supply funciton for a Fermi-degenerated system. The upper shows a schematic example of transimission coefficient $\mathcal{T}$ which has a peak at $E_r - eV/2$. Right panel: Schematic drawing of current-voltage characteristics expected from the left supply function and transmission coefficient.

eV exceeds E_F. For simplicity, we use this form to consider the I-V characteristics. The zero energy is now taken to the band edge of source electrode and let the resonant energy level in the well be E_r then $\mathcal{T}$ has a peak at $E_r - eV/2$. Hence resonant tunneling current appears with the threshold voltage around $E_r - eV/2 \sim E_F$ as illustrated in the right panel of Fig.3.19 and rapidly goes down to zero at $E_r - eV/2 \sim 0$. Actual current contains thermal excitation, incoherent tunneling, etc. and very roughly the lineshape sketched as the broken line is expected.