We have seen the effect of confinement with heterojunctions. When the barrier width is finite, the transport across the barrier with quantum tunneling. These can be viewed as elementary quantum mechanics though it is an important step that such phenomena can be observed in real systems in the very beginning of semiconductor quantum physics. Actually the heterojunction technique leads to the prosperity of the field and many novel devices have been created.

7.1.3 Excitons in two-dimensional systems

In the previous section, in the absorption spectrum of a quantum well, we observed peak structures around the absorption edges. They are from the excitons explained in Sec. 3.3.2 thought the lowering of the spatial dimension results in some quantitative differences from the excitons in the bulk. We would like to have a brief look at the excitons in both two and three dimensions.

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

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 + V_{\mathrm{c}}(\boldsymbol{r})\right)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{7.9}$$

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

$$V_{\mathrm{c}}^{\mathrm{2d}}(\boldsymbol{r}) = -\frac{e^2}{4\pi\epsilon\epsilon_0|\boldsymbol{r}|}, \qquad V_{\mathrm{c}}^{\mathrm{1d}}(r) = -\frac{e^2}{4\pi\epsilon\epsilon_0(|z| + 0.3r_0)}, \tag{7.10}$$

particularly for one-dimensional (along z-axis) systems. This is because simple transformation of eq.(7.9) into one-dimension causes anomalous behavior including divergence of binding energy. The potential form in eq.(7.10) 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.(7.9) can be written in the forms

$$\psi^{\mathrm{3d}} = \rho^l e^{-\rho/2}R(\rho)Y_{l,m}(\theta, \varphi), \quad \psi^{\mathrm{2d}} = \rho^{|m|}e^{-\rho/2}R(\rho)e^{im\varphi}, \quad \psi^{\mathrm{1d}} = R(\zeta). \tag{7.11}$$

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

$$\rho = \alpha r, \quad \zeta = \alpha(|z| + 0.3r_0), \quad \alpha = \frac{\sqrt{-8m^*E}}{\hbar}. \tag{7.12}$$

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

$$\begin{cases} \left(\rho\dfrac{\partial^2}{\partial\rho^2} + (p+1-\rho)\dfrac{\partial}{\partial\rho} + q\right)R(\rho) = 0 : & \text{3-, 2-dimensional,} \\[3mm] \left(\dfrac{\partial^2}{\partial\zeta^2} + \dfrac{\partial}{\partial\zeta} + \dfrac{\lambda}{\zeta}\right)R(\zeta) = 0, \;\; \lambda \equiv \dfrac{e^2}{4\pi\epsilon_0\hbar}\sqrt{-\dfrac{m^*}{2E}} : & \text{1-dimensional,} \end{cases} \tag{7.13}$$

where p, q are

$$p = \begin{cases} 2l+1 & \text{(3-dimensional)} \\ 2|m| & \text{(2-dimensional)} \end{cases}, \qquad q = \begin{cases} \lambda - l - 1 & \text{(3-dimensional)} \\ \lambda - |m| - 1/2 & \text{(2-dimensional)} \end{cases}, \tag{7.14}$$

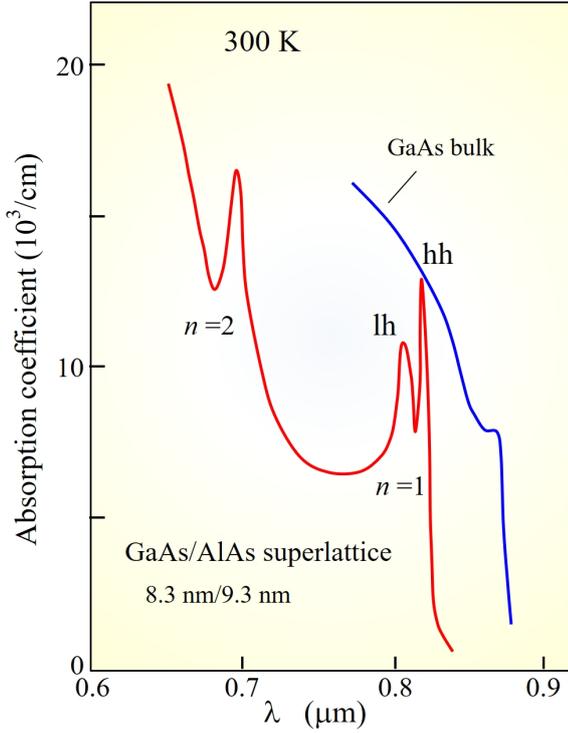where l is angular momentum quantum number and m is magnetic quantum number.

Fig. 7.3 Exciton absorption peaks appeared in the absorption spectrum of a GaAs(8.3 nm)/AlAs(9.3 nm) superlattice at room temperatre (red line). The coupling between the quantum well in the superlattice is weak. n in the figure is the subband index and different from the one of excitonic states index. The peak at the ground subband ($n = 1$) shows the splitting into lh and hh[1]. The $n = 2$ peak is considered as from hh reduced mass.

For three and two dimensional systems, $R(\rho)$ in eq.(7.13) is expanded as follows.

$$R(\rho) = \sum_\nu \beta_\nu \rho^\nu, \quad \beta_{\nu+1} = \beta_\nu \frac{\nu - q}{(\nu + 1)(\nu + p + 1)}. \tag{7.15}$$

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.

$$n \equiv \lambda = \nu_{\max} + l + 1 \text{ (3-dimensional)}, \quad n \equiv \lambda - \frac{1}{2} = \nu_{\max} + |m| \text{ (2-dimensional)}. \tag{7.16}$$

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

$$E_{bn}^{3d} = -\frac{E_0}{n^2} \qquad\qquad n = 1, 2, \cdots, \tag{7.17}$$

$$E_{bn}^{2d} = -\frac{E_0}{(n + 1/2)^2} \qquad\qquad n = 0, 1, \cdots. \tag{7.18}$$

Here the energy unit E_0 is

$$E_0 = \frac{e^2}{8\pi\epsilon\epsilon_0 a_0^*}, \quad a_0^* = \frac{4\pi\epsilon\epsilon_0 \hbar^2}{m^* e^2}, \tag{7.19}$$

where a_0^* is the effective Bohr radius. From eq.(7.16), 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 $\psi_{1s}^{2d} \propto \exp(-r/a_0^{*2d})$, (7.13) ⌢ $l = m = 0$ 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.

E9-2

In Fig. 7.3, we show the absorption coefficient of a GaAs(8.3 nm)/AlAs(9.3 nm) superlattice (red line) and taht of a high purity bulk GaAs (blue line). In the bulk line, a shoulder structure at the absorption edge is observed. On the other hand, the exciton absorptions take clear peak structures. Furthermore, the peak at the absorption edge of 1st subband ($n = 1$) shows a clear splitting due to two effective masses lh and hh in the valence band, which result in the two reduced masses. The exciton peak at the second subband edge is also clearly observed and considered as from hh reduced mass. These observations are possible by the above enhancement in the binding energy due to the confinement.

7.2   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**.

7.2.1   Transfer matrix

Let us consider a region Q in a one-dimensional space and as shown in Fig. 7.5(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. We define $\kappa \equiv \sqrt{2mV_0}/\hbar$. 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

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

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

$$A_1 + B_1 = V_1 + W_1, \qquad A_2 + B_2 = e^{-\kappa L} V_1 + e^{\kappa L} W_1, \tag{7.20}$$

$$ik(A_1 - B_1) = \kappa(-V_1 - W_1), \qquad ik(A_2 - B_2) = \kappa(-e^{-\kappa L} V_1 + e^{\kappa L} W_1). \tag{7.21}$$

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

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} \equiv M_T \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{7.22}$$
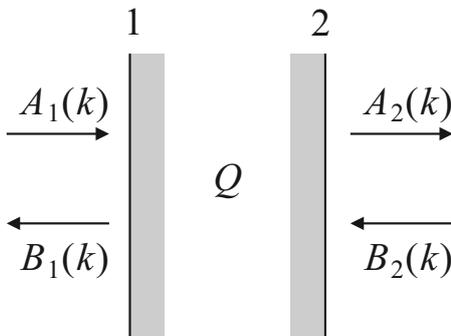

Fig. 7.4   Scheme of T-matrix

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

$$
\begin{cases}
m_{11} = \left[\cosh(\kappa L) + i\dfrac{k^2 - \kappa^2}{2k\kappa}\sinh(\kappa L)\right], \\[2mm]
m_{12} = -i\dfrac{k^2 + \kappa^2}{2k\kappa}\sinh(\kappa L), \\[2mm]
m_{21} = m_{12}^*, \qquad m_{22} = m_{11}^*.
\end{cases}
\tag{7.23}
$$

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 (7.22) guaranteed by the linearity of Schrödinger equation. A matrix like M_T is called **transfer matrix** (T-matrix).

In Eq.(7.23), 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 (7.22), (7.23) as

$$
t \equiv \frac{A_2}{A_1} = \frac{|m_{11}|^2 - |m_{12}|^2}{m_{11}^*} = \frac{1}{m_{11}^*} = \frac{2ik\kappa}{(k^2 - \kappa^2)\sinh(\kappa L) + 2ik\kappa\cosh(\kappa L)},
\tag{7.24}
$$

$$
r \equiv \frac{B_1}{A_1} = -\frac{m_{21}}{m_{22}} = \frac{(k^2 + \kappa^2)\sinh(\kappa L)}{(k^2 - \kappa^2)\sinh(\kappa L) - 2ik\kappa\cosh(\kappa L)}.
\tag{7.25}
$$

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

$$
\text{Transmission: } T = |t|^2, \qquad \text{Reflection: } R = |r|^2, \qquad |t|^2 + |r|^2 = 1,
\tag{7.26}
$$

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

$$
M_T = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix}.
\tag{7.27}
$$

7.2.2 Transmission through double-barrier structure

Let us consider the transmission through the double barrier potential illustrated in Fig. 7.5. 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 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 (7.23) 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

$$
M_W = \begin{pmatrix} \exp(ikW) & 0 \\ 0 & \exp(-ikW) \end{pmatrix}.
\tag{7.28}
$$

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.
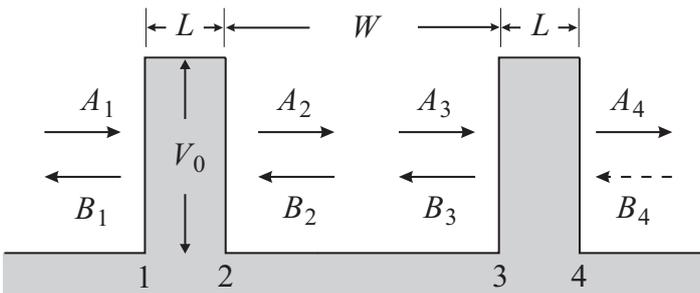

Fig. 7.5 Schematic illustration of double barrier potential.

E9-4

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

$$M_{DW} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} e^{ikW} & 0 \\ 0 & e^{-ikW} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \equiv \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}. \tag{7.29}$$

The transmission coefficient is, from (7.29),

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

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

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

The the transmission coefficient is obtained as

$$T = \frac{1}{|T_{11}|^2} = \frac{1}{1 + 4|m_{11}|^2 |m_{12}|^2 \cos^2(\varphi + kW)}. \tag{7.30}$$

The final form of transmission coefficient is then in combination obtained with (7.23).

Figure 7.6(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

$$\varphi + kW = \left(n - \frac{1}{2}\right)\pi \quad (n = 1, 2, \cdots), \tag{7.31}$$

from (7.30), where φ is witten from (7.23) as

$$\varphi = \arctan\left[\frac{k^2 - \kappa^2}{2k\kappa} \tanh(\kappa L)\right], \tag{7.32}$$

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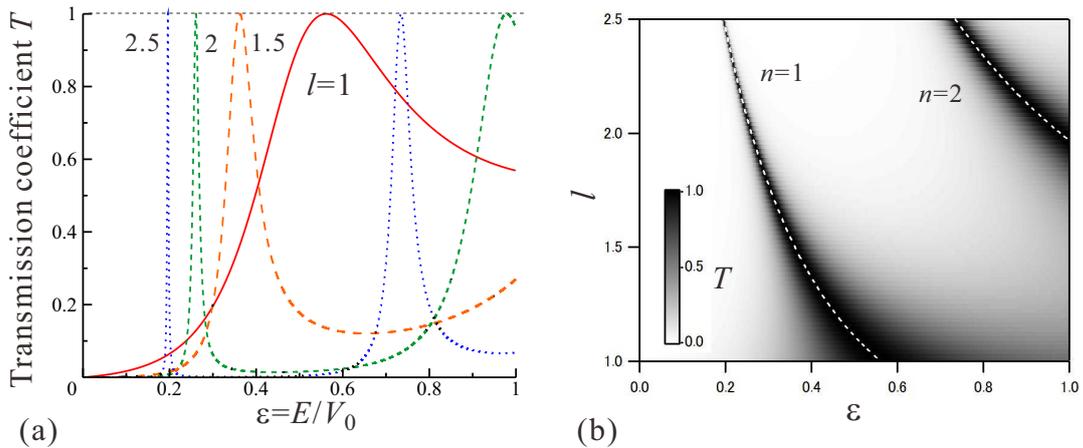

(a) (b)

Fig. 7.6 (a) Transmission coefficient T calculated on (7.30) 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 (7.31), (7.32).

In Fig. 7.6(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.

7.2.3 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 ΔE_v =0.47 eV. There are two species of holes at the top of valence band in GaAs with effective masses $0.51 m_0$ and $0.082 m_0$, 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. 7.7(a), widths of 5nm both for the barriers and the well. The barriers and the well parts do not have any doping. Figure 7.7(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. 7.7 is displayed as a function of energy in Fig. 7.8. 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 7.7(a) shows the positions of resonance levels in the well numerically calculated from eq.(7.31).

In order to see the behavior of tunneling, usually source-drain voltage V_sd is applied as illustrated in Fig. 7.7(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.
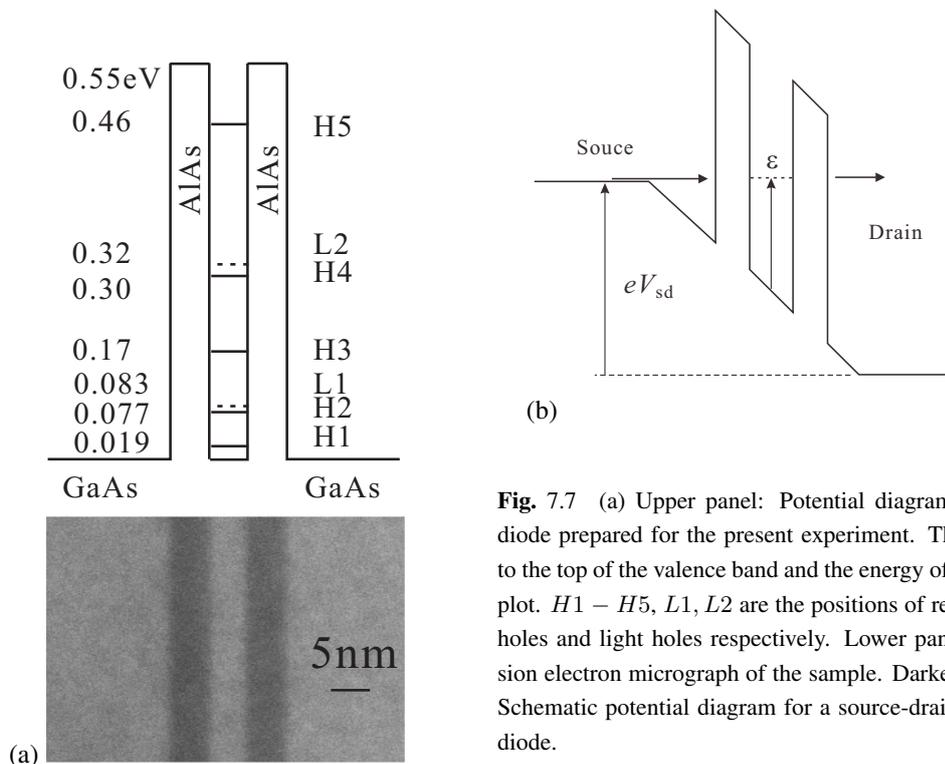

(a)

(b)

Fig. 7.7 (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.
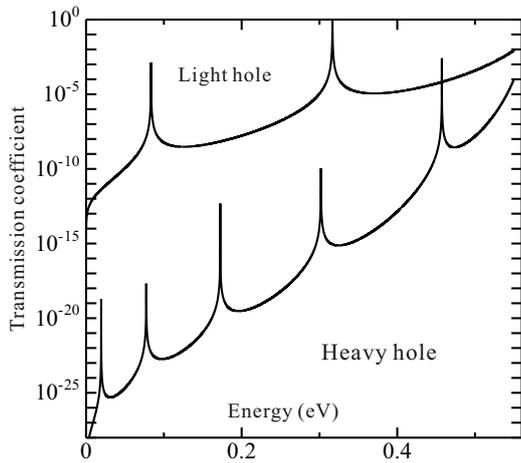
Fig. 7.8 Energy dependence of transmission coefficient for the double barrier structure with parameters given in the text and with eq.(7.30). The peaks hit 1 actually but too narrow to be sampled.

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 measured current-voltage curve in a double barrier diode (the one in Fig. 7.7) is shown in Fig. 7.9(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. 7.9(b) [*1].
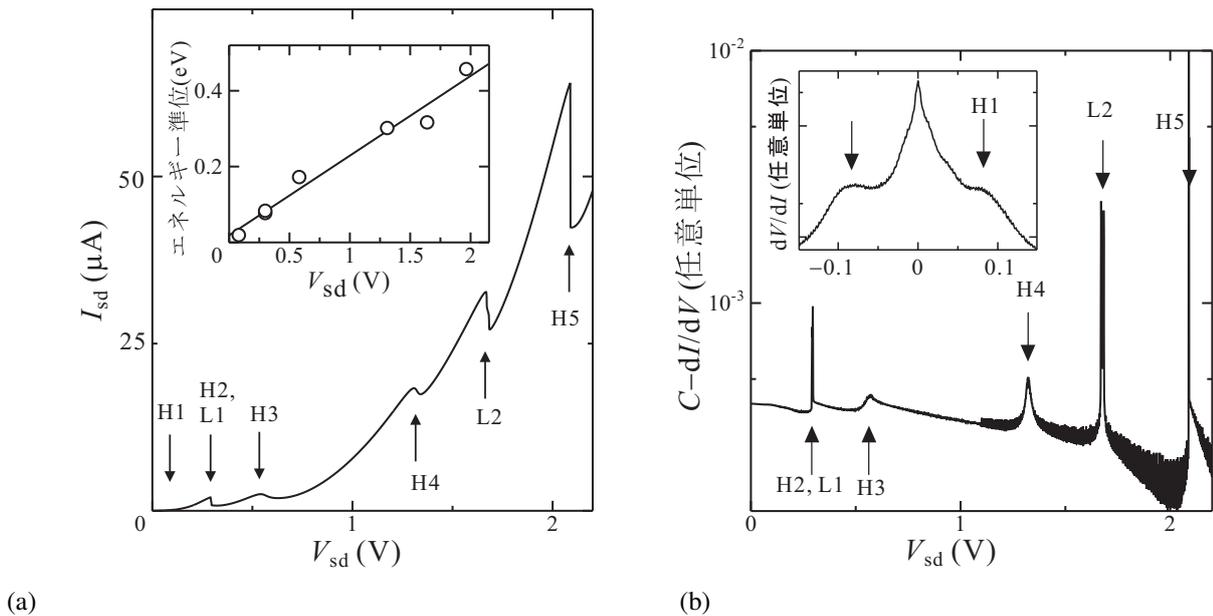

(a)

(b)

Fig. 7.9 (a) Current-voltage characteristics of the double barrier diode introduced in Fig. 7.7. 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.

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

7.2.4 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. 7.10. 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 $V_{\mathrm{KP}}(x)$ and write down the Schrödinger equation as

$$\left[-\frac{\hbar^2 d^2}{2m dx^2} + V_{\mathrm{KP}}(x)\right]\psi(x) = E\psi(x), \quad V_{\mathrm{KP}}(x) = V_{\mathrm{KP}}(x+d). \tag{7.33}$$

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.

$$\psi_K(x) = u_K(x)e^{iKx}, \quad u_K(x+d) = u_K(x), \quad K \equiv \frac{\pi s}{Nd}. \tag{7.34}$$

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

$$M_d(k) = \begin{pmatrix} e^{ikW} & 0 \\ 0 & e^{-ikW} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} m_{11}e^{ikW} & m_{12}e^{ikW} \\ m_{21}e^{-ikW} & m_{22}e^{-ikW} \end{pmatrix}. \tag{7.35}$$

As before, we write the input/output in the left hand side of i-th cell as (a_i, b_i), then from (7.34),

$$\begin{pmatrix} a_{i+1} \\ b_{i+1} \end{pmatrix} = M_d \begin{pmatrix} a_i \\ b_i \end{pmatrix} = e^{iKd} \begin{pmatrix} a_i \\ b_i \end{pmatrix} \tag{7.36}$$

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 (7.36), the two eigenvalues $e^{\pm iKd}$ are obtained. We re-use $\{m_{ij}\}$ in (7.23) to get to the equations

$$e^{iKd} + e^{-iKd} = 2\cos Kd = \mathrm{Tr}M_d = 2\mathrm{Re}(e^{-ikW}m_{11}^*), \tag{7.37}$$

$$\cos[K(L+W)] = \cosh(\kappa L)\cos(kW) - \frac{k^2 - \kappa^2}{2k\kappa}\sinh(\kappa L)\sin(kW). \tag{7.38}$$

By use of φ in (7.32), expression

$$\cos(Kd) = |m_{11}|\cos(kW + \varphi) = \frac{1}{|t|}\cos(kW + \varphi) \tag{7.39}$$

is available.

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

$$\cos(Kd) = \cos(kd) + \frac{mC}{\hbar^2 k}\sin(kd). \tag{7.40}$$
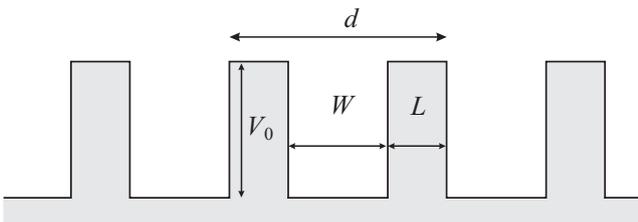

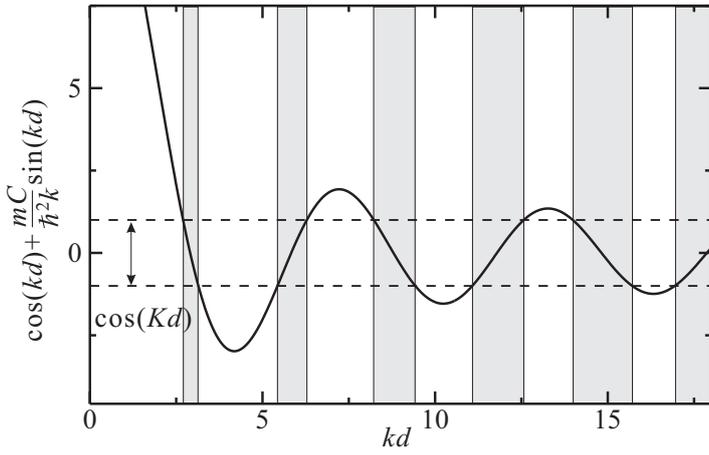**Fig.** 7.10 One dimensional rectangular potential (Kronig-Penny type potential)

Fig. 7.11 RHS of (7.40) as a function of kd. Here $mdC/\hbar^2$ is taken to be 13. The gray belts indicate "allowed bands".

Figure 7.11 shows the RHS as a function of kd. The solution K for (7.40) 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

$$E(K) = \frac{E_{nw}}{2}(1 - \cos Kd). \tag{7.41}$$

The group velocity and the effective mass are

$$v_g(K) = \frac{E_{nw}d}{2\hbar}\sin Kd, \quad m^*(K) = \frac{\hbar^2}{E_{nw}d^2}\sec Kd. \tag{7.42}$$

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

$$m^*\frac{dv}{dt} = \hbar\frac{dK}{dt} = F = eE_m. \tag{7.43}$$

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

An acceleration according to (7.43) 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 (7.42),

$$v_g(t) = \frac{E_{nw}d}{2\hbar}\sin\left(\frac{eE_m d}{\hbar}t\right), \quad x(t) = \frac{E_{nw}}{2eE_m}\left[1 - \cos\left(\frac{eE_m d}{\hbar}t\right)\right]. \tag{7.44}$$

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.

7.3 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. 7.12, 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
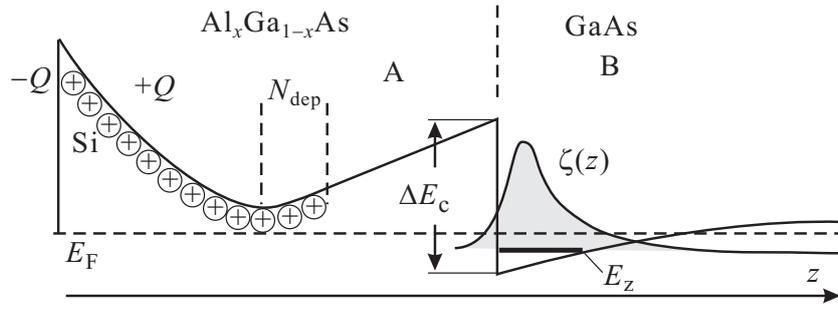
Fig. 7.12 Schematic cross sectional view of two-dimensional electrons at a modulation doped $Al_xGa_{1-x}As$/GaAs heterointerface.

approximation[*2]. 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.

z-axis is taken to be perpendicular to the heterointerface plane. As in Fig. 7.12, 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_{ddep}. The electrostatic potential from the charges is, far inside the lattice from the doping region, $V_D(z) = (4\pi e^2/\epsilon\epsilon_0)N_{dep}z$. 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

$$V(z) = V_h(z) + \frac{4\pi e^2}{\epsilon\epsilon_0}\left[N_{dep}z - n_{2d}(E_z)\int_{-\xi}^{\infty}|z-z'||\zeta(z')|^2dz'\right]. \tag{7.45}$$

Schrödinger equation for $\zeta(z)$

$$\left[-\frac{\hbar^2}{2m^*(z)}\frac{\partial^2}{\partial z^2} + V(z)\right]\zeta(z) = E_z\zeta(z) \tag{7.46}$$

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

$$\zeta(0)^{(A)} = \zeta(0)^{(B)}, \quad \frac{1}{m_A^*}\frac{d\zeta^{(A)}}{dz}\bigg| = \frac{1}{m_B^*}\frac{d\zeta^{(B)}}{dz}\bigg|. \tag{7.47}$$

In the Poisson-Schrödinger procedure, one should solve the equations from (7.45) to (7.47) 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.

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

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

$$\zeta(z) = \sqrt{\frac{b^2}{2}} z \exp\left(-\frac{bz}{2}\right) \tag{7.48}$$

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

$$b^3 = \frac{48\pi me^2}{\epsilon\epsilon_0\hbar^2}\left(\frac{11}{32}n_{2d} + N_d\right). \tag{7.49}$$

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.[3].

7.4   Fabrication of quantum wires

Nowadays we have so many methods to fabricate quantum wires and reviewing in this narrow space is impossible. Here we have a short look at a few examples of them.

7.4.1   Split gates, other physical approach

The split gate method starts from a 2DEG. Metallic films on the surface form Schottky barriers and deplete the electrons underneath them. Then we can build potentials with various shapes through those of the metals.

In the split gate method, enlargement of depleted regions with reverse (negative) bias voltage V as in Fig. 7.13(a) is often used. Let us consider a simple model illustrated in Fig. 7.13(b), where two half-infinite metals are placed with a distance w. The line density of charge σ, created by applied gate voltage is assumed to be uniform. The electric field formed by these charges has the z-component $\mathcal{E}_z(x, d)$ as,

$$\mathcal{E}_z(x, d) = \frac{-\sigma}{2\pi\epsilon\epsilon_0}\left[\pi + \arctan\left(\frac{x - w/2}{d}\right) - \arctan\left(\frac{x + w/2}{d}\right)\right]. \tag{7.50}$$

(7.50) depends on d, but as a coarse approximation, we ignore the dependence within the depth η of the two-dimensional electron gas (2DEG) potential and the potential modulation due to the split gate is summarized as $V_{\text{sg}}(x) = e\eta\mathcal{E}_z(x, \eta)$.
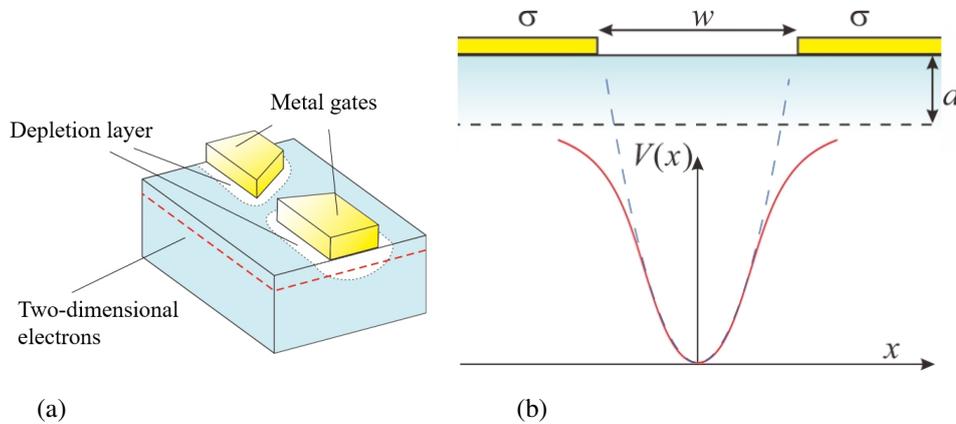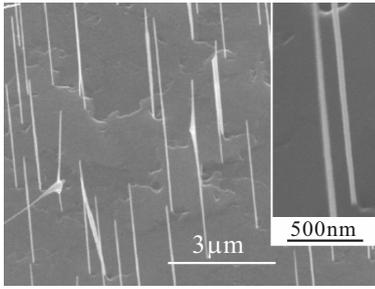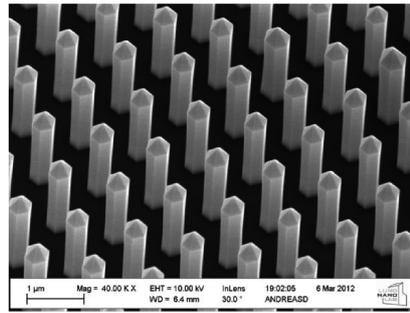

Fig. 7.13   (a) Schematic illustration of micro-fabrication by split-gate method. (b) Electrostatic potential formed by a split gate calculated on a simple model in the text.

Fig. 7.14 (a) Scanning electron micrograh of InAs nanowires along [111] grown by vapor-liquid-solid phase method with metal particles as catalyst on an InAs(111)B substrate. (b) GaN-InGaN core-shell type nanowires.

An example of $V_{sg}(x)$ is shown in the lower panel in Fig. 7.13(b). As indicated by the broken line, a parabollic potential well approximates $V_{sg}(x)$ around the bottom [*3].

The gate electrodes in Fig. 7.13(b) hence give confinement along x in addition to z-direction heterinterface confinement potential. The kinetic energies for these two directions are quantized and quantum wires are realized. Density of states in one-dimensional systems has a divergence at the bottom of band, which corresponds to the descrete energy levels in the potential in Fig. 7.13(b).

Another way to form quantum wires with physical methods is "cutting" of 2DEG into thin and long shapes with wet (*i.e.* in some solvent) or dry (*i.e.* in some plasmas) etching method. There are also several ways such as regrowth of heterostructures onto a cleaved edge of another heterostructure to form T-shaped thin-line potential, ion-inplantation inactivation, etc.

7.4.2   Self-assembled nanowires

After putting some "seeds" onto semiconductor substrates by electron beam irradiation, etc., crystal growth onto it causes nanowire growth at the seeds under some conditions. In such a growth, with changing "flying" materials heterostructures or doping can be installed in the wire. Figure 7.14(a) shows InAs nanowire along [111] grown on an InAs(111)B substrate with Au nanoparticles as the catalyst by vapor-liquid-solid phase method. Fiugre 7.14(b) shows GaN-InGaN core-shell type nanowires.

Well-known **carbon nanotubes**, which are rolled up graphenes with nanometer-size diameters, may not be classified into one-dimensional systems, are also a kind of "self-assembled systems".

7.5   Fabrication of quantum dots

Fabrication methods of quantum dots are also classified into physical methods, self-assemble methods, and their combinations. I would like to introduce just a part of them.

7.5.1   Physical method

Quantum dots (QDs) are expected to have a very wide range of optical applications, and quantum dot lasers are already on the market. Here, however, we restrict ourselves to the QD for the study of transport. To measure the electric conduction, we need to touch electrodes to the dots. As the electrode material we consider a normal metal with an ordinary Fermi surface. And as the "connection", we consider tunneling junctions, through which electron can transmit

---

[*3] Anyway a rounded bottom of symmetric potential can often be approximated by a parabolla because the leading term in the power expansion is usually of the second order.

with quantum tunneling. At least a single electorode is required. And to compose structure for measurement of transport between two particle reservoirs just like quantum wires, FET, two electrodes, hence two tunneing junctions are required. As shown in Fig. 7.15, these electrodes are called just like FET, source and drain.

In quantum dots, the density of states is like a series of δ-functions. The electric conduction is determined by the tunneling probability of junctions and the relative positions of the δ-function like density of states and the Fermi levels in the drain and the source. If we place another metallic electrode close to the dot without tunneling probability, the electrochemical potential of the dot can be controlled with the electric field from the electrode. The electrode is called gate. Figure 7.15(a) shows the total schematic of the quantum dot for conduction measurement.

The split-gate method introduced for the quantum wire can also be applied to form quantum dots. Figure 7.15(b) illustrates a possible pattern of Schottky electrodes. For the tunnel junctions, quantum point contacts near the pinch-off condition are used. For the gates, also Schottky electrodes are used in the reverse bias condition. As can be imagined from the figure, because the reverse bias voltage enlarges the depletion layer to squeeze the dot region, the size of the quantum confinement potential is diminished that widens the level intervals other than the single electron effect, which will be discussed later. Hence the gate is sometimes called "plunger" gate. In this kind of configuration, the source, dot, drain line up side by side along the two-dimensional electrons, hence called "lateral" quantum dots. When the number of electrons in a quantum dot is reduced by the gate voltage, the dot size also becomes smaller, it is spatially separated from the source and drain, the tunnel probability becomes smaller, and conduction becomes unmeasurable. This once considered as a difficult problem but has been overcome by the remote charge detection. The structure in Fig. 7.15(c) is made from the double barrier structure The structure shown in Fig. 7.15(c) is made by cutting out a double barrier structure into a cylinder shape. Then Schottky gate electrodes are deposited on the side of the cylinder. This is called a "vertical" quantum dot. The tunnel couplings are determined by the double barrier structure, not affected by the electron number. The property makes the structure apropriate for the experiments for small number of electrons. There is a problem in connection with external quantum circuits, which requires various devicing.

7.5.2 Self assembling method

The epitaxial growth has various "mode" in the growth process. The layer-by-layer growth mode is called Frank-van der Merve (FvdM) mode (Fig. 7.16(a)). When the interface energy accumulation between the thin film and the substrate is large due to the combination of materials, the deposited material is repelled from the substrate in the beginning of the growth and a three dimensional growth begins. As a such growth mode, the Volmer-Weber (VM) mode is illustrated in Fig. 7.16(b). In Fig. 7.16(c) we show the Stranski-Krastanow (SK) mode, in which the growth is two-dimensional at the very beginning but changes into three-dimensional due to the lattice distortion inside the film. With such three
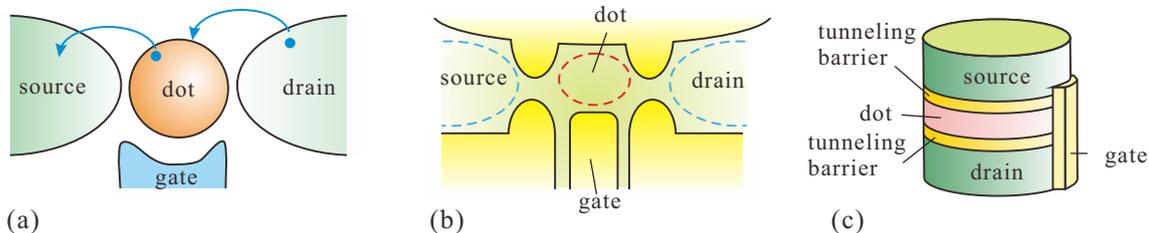

(a)   (b)   (c)

Fig. 7.15   (a) Schematic of quantum dot structures for transport measurement. Two electrodes for examining conduction, a source, and a drain are connected via a tunnel barrier across the quantum dot, and a gate electrode that controls the potential of the dot is arranged at a distance. (b) Illustration of a "lateral" quantum dot. Nano-fabricated metallic gates on two-dimensional electrons are used. (c) Illustration of a "vertical" quantum dot. The dot layer is between two barrier layers and the doped upper and lower layers are the source and drain. The layers are cut to a piller and the metallic gate is deposited surrounding it (in the figure a part of the gate is drawn).

dimensional growth, the structures with dimensions less than two can be obtained.
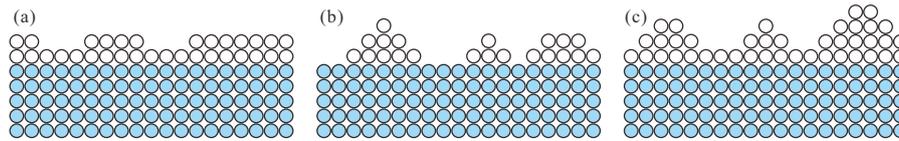

Fig. 7.16   Various epitaxial growth modes. Blue-gray circles represent atoms in the substrate, while open circles do the atoms in the fim. (a) Frank-van der Merve. (b) Volmer-Weber. (c) Stranski-Krastanow.

A representative example of self-assembling of low-dimensional systems is the growth of quantum dots with SK mode growth. Such examples are shown in Fig. 7.17. When some amount of a semiconductor InAs, which has the lattice constant 7% larger than that of GaAs, is deposited on a GaAs substrate and kept at some high substrate temperature for some time, the indium atoms on the substrate migrate and accumulate to form quantum dot structures.

In the case of InAs, dots are self-assembled because when a relatively small number of In atoms are present on the substrate, it is more energetically stable to perform three-dimensional growth to escape from the interface, which gets strong lattice distortion from the substrate. The indium atoms first form a two-dimensional wetting layer with thickness of a few lattice constant then dots are randomly formed in the shape that depends on the crystal direction of the surface. The quantum dots produced by the SK mode growth have random sizes and positions. On the other hand, the SK dots are with high crystal qualities and with high densities, and thus widely used for optical devices like quantum dot lasers. In addition, since InAs has low junction resistances with metals, conduction measurement is also performed by attaching metal electrodes. A method often adopted as a combination of self-assembling and physical methods is to attach gates and barrier electrodes to self-forming nanowires by lithography to make dots. In particular, the self-assembling of InSb or InAs, to which heterojunction technique is difficult (though not impossible) to apply, is used to form quantum dots and other structures with many gate electrodes. There is also a method to form quantum dots with implementing the barrier layers into nanowire during the growths.
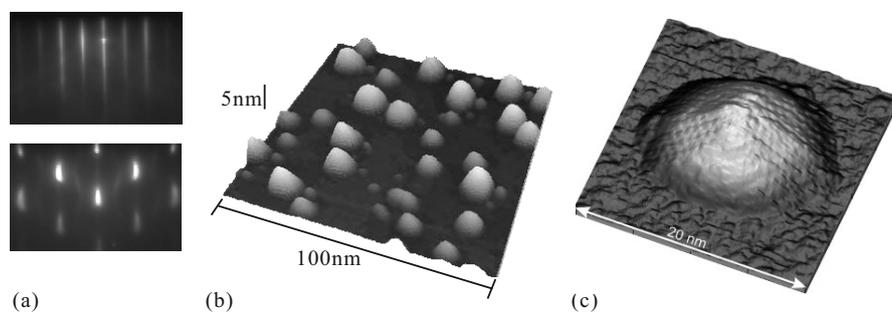

Fig. 7.17   (a) InAs quantum dot growth on GaAs (001) substrate. Upper panel: RHEED pattern of two-dimensional growth at the beginning of InAs growth. Lower penel: RHEED pattern of three-dimensional growth of quantum dot structure. (b) Atomic force micrograph of self-assembled quantum dots. (c) Scanning tunneling micrograph of a quantum dot. The lattice image can be seen, manifesting that the whole dot is a single crystal.

7.5.3   Colloidal quantum dots

In recent years, the colloidal manufacturing method has come to be widely used to form optical quantum dots for optical use. As shown in Fig. 7.18, this is a method of obtaining quantum dots by injecting a dot material called a "precursor" into a solvent, dissolving it, making it supersaturated by a temperature change, and precipitating a part of it. From the relationship between surface area and volume, when the degree of supersaturation falls below a certain value, dots that continue to grow and dots that redissolve are separated, so dots with relatively uniform sizes can be obtained. This is

called Ostwald reipening. After the growth reached saturation, surface covering of the grown dots with another material is possible by adding a different precursor into the solvent. Such covered quantum dots obtained in this way are called **core-shell** type quantum dots. The luminescence wavelength of the quantum dots can be tuned by their size and hence it is possible to form high efficiency luminous materials. They are already applied to, *e.g.* quantum dot displays.
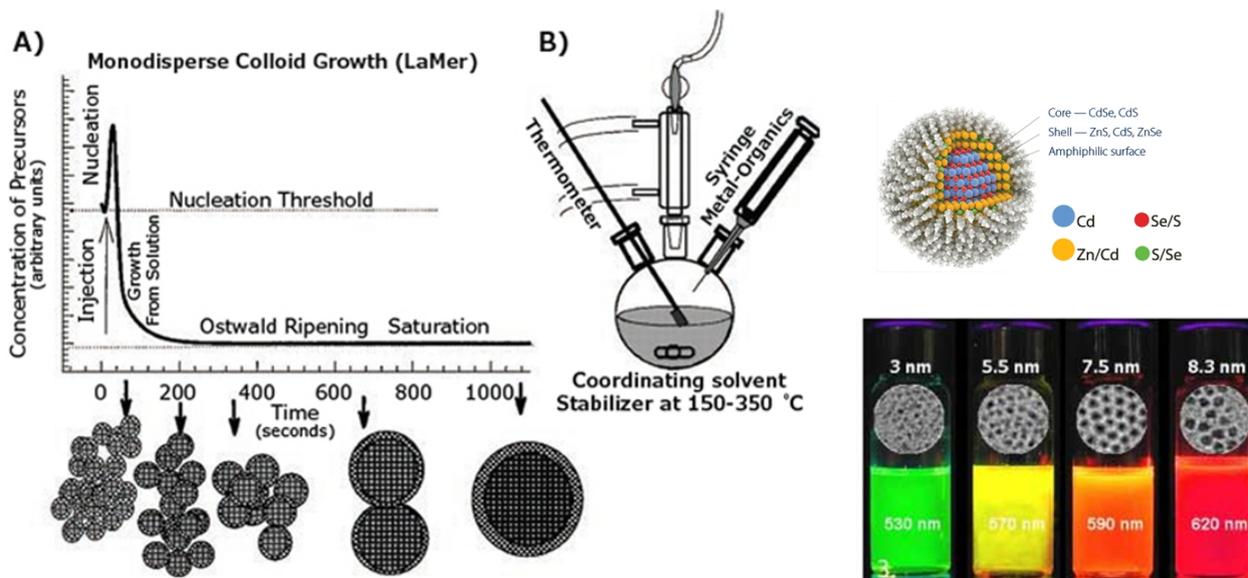

Fig. 7.18 Left (A): Illustration of quantum dot formation with precipitation from supersaturated solvent. The time dependence of the density of precursor is plotted in the graph. Left (B): Illustration of experimental setup of precipitation method. Right upper: Schematics of core-shell type quantum dots. [4] Right lower: Illumination from CdS-based core-shell quantum dots and TEM images. (Ocean Nanotech. web site. https://www.oceannanotech.com/)

References

[1] T. Ishibashi，S. Tarucha and H. Okamoto, Proc. GaAs and Related Compounds, p. 587 (Oiso, 1981).

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

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

[4] C. B. Murray, C. R. Kagan, and M. G. Bawendi, Annu. Rev. Mater. Sci. **30**, 545 (2000).