## 8.7 Single electron effect and quantum confinement

In the end of the chapter, we will have a look on single electron effect and quantum confinement to zero-dimensional system in quantum dots.

### 8.7.1 Single electron effect

In the transport through quantum dots the first importantce is on the single electron effect. The single electron effect is in very short, the the electrostatic energy of an electrostatically isolated system changes with adding(extracting) an electron, and when this increase is larger than the thermal fluctuation, the tunneling of the electron is prohibited. This effect is called **Coulomb blockade**. The electrostatic energy of a quantum dot is described with a capacitance $C$ of the dot and the electrostatic energy of charging by a single electron is $E_c = e^2/2C$, which is finite and even can be large for small $C$. As a first approximation we separates the electronic states into two: states inside the dot and those outside the dot. Hence the number of electrons in the dot takes an integer (descrete value). There are two possible simplest transport processes of single electrons from a source to a drain: the dot catches an electron from the source then releases one to the drain, and conversely the dot releases one to the drain first, then catches one from the source.

Let us take the simplest **constant interaction** model, in which any pair of electrons in the dot has the same (constant) Coulomb interaction energy $U$. Then the total Coulomb energy in the dot with $N$ electrons is

$$E_{cN} = {}_NC_2U = \frac{N(N-1)U}{2} = \frac{U(N-1/2)^2}{2} - \frac{U}{8}. \tag{8.59}$$

The variation in the Coulomb energy with the transition $N \to N+1$ is

$$\Delta E_+(N) = (N-1)U. \tag{8.60}$$

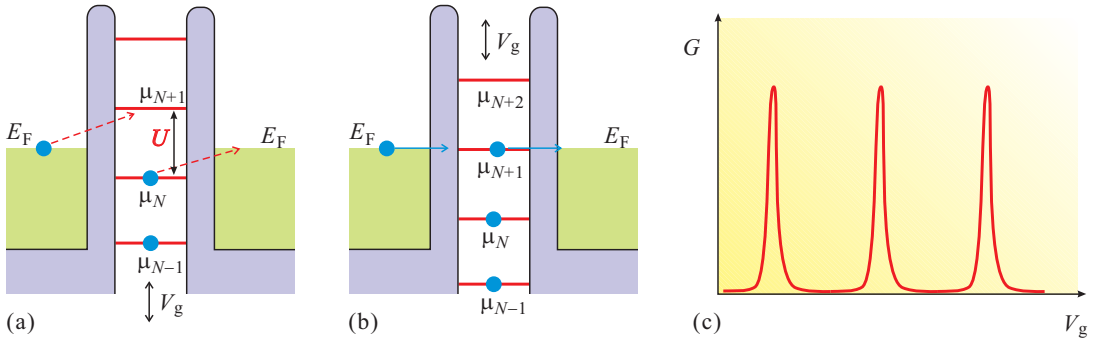



**Fig.** 8.26 (a) Schematic diagram of chemical potentials in the constant interaction model, in which the chemical potentials are descrete with the same distance $U$ (the amplitude of two-electron interaction). At zero-bias, when none of the descrete chemical potentials meets the Fermi level $E_F$ in the source and drain, a finite energy is required for an electron to tunnel, which prohibits the tunneling (Coulomb blockade). (b) When the origin of the descrete chemical potentials is shifted by the gate voltage $V_g$ and one of them hits $E_F$, the tunneling thus the electric conduction becomes possible. (c) When $V_g$ is swept, a repetition of processes (a) and (b) results in the series of sharp peaks with a regular interval in the quantum dot conductance $G$ (Coulomb oscillation).

If we ignore other kinds of energy, $\Delta E_+(N)$ should be the electrochemical potential of $N$-th electron and from Eq./ (8.60) we see that the electrochemical potentials are ordered with the same distance being proportional to $N$.

### 8.7.2 Coulomb oscillation, Coulomb diamond

Let us write the electrochemical potential of the dot with $N$-electrons $\mu_N$, and let $N_0$ be the electron number when the dot is in equilibrium with the electrodes, then $\mu_{N_0} < E_{\rm F} < \mu_{N_0+1}$. When $\mu_{N_0}$ is equal to $E_{\rm F}$ (Fermi energy in the electrodes), electrons can go into or out from the dot with tunneling from the electrodes, then at zero source-drain voltage ($V_{\rm sd} = 0$) the electric conductance ($G(0)$) takes a finite value. When $\mu_{N_0}$ does not hit $E_{\rm F}$, the tunnleing of an electron between the electrodes and the dot requires a finite energy and is prohibited (Coulomb blockade). As in Fig. 8.26(a), (b), that condition of finite $G(0)$ appears with a constant interval. Hence $G(0)$ forms regular peaks for a sweep of $V_{\rm g}$ as shown in Fig. 8.26(c), which is called **Coulomb oscillation**.

The constant interaction model can also be described as a simple circuit model illustrated in Fig. 8.27(a). Here the charge of an electron is $-e$.

$$Q_1 + Q_2 = -eN, \quad Q_1 = CV_{\rm d}, \quad Q_2 = C_{\rm g}(V_{\rm d} - V_{\rm g}), \tag{8.61}$$

and the charging energy is

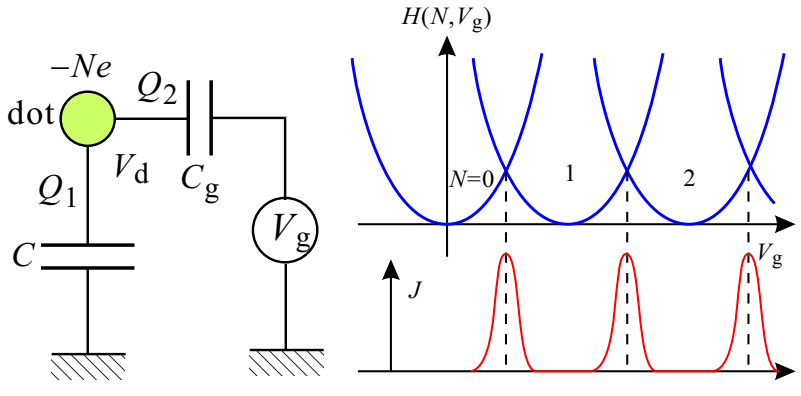
$$E = \frac{1}{2}CV_{\rm d}^2 + \frac{1}{2}C_{\rm g}(V_{\rm d} - V_{\rm g})^2, \tag{8.62}$$

in which the second term is the integral of the work done by the power source connected to the gate electrode from voltage 0 to $V_{\rm g}$. When we thermodynamically treat the problem whether the process proceeds or not under the condition that some system outside automatically provides energy, we need to consider **enthalpy**, which in the case of the pressure of atomosphere, written as $H = U - PV$. Here $PV$, the product of pressure and volume, corresponds to the automatic energy supply corresponding to the second term in Eq. (8.62). Then from (8.61) and (8.62),

$$H(N, V_{\rm g}) = \frac{(Ne - C_{\rm g}V_{\rm g})^2}{2(C + C_{\rm g})}. \tag{8.63}$$

If we plot this as a function of $V_{\rm g}$, as shown in Fig. 8.27(b), parabollas are lined up corresponding to $N$ and the Coulomb peaks appear at the crossing points of the prabollas.

Next we consider the case that the gate voltage is fixed, the drain is grounded, and the source voltage is swept. The simplest model for such situation is shown in Fig. 8.28. At the positions of Coulomb peaks, the topmost chemical potential of the dot hits the Fermi level in the source and drain electrodes. The number of electrons differs by 1 at (a) and at (c). When $V_{\rm g}$ is at (b), the current is blocked at zero bias. Then if we increase the source voltage (decrease $E_{\rm F}$ in the source), and the chemical potential position used for the conduction at (a) goes in between the Fermi levels of source




(a) (b)

**Fig.** 8.27 (a) Simple model of the single electron charging in a quantum dot described with a self capacitance $C$ and a gate capacitance $C_{\rm g}$. (b) Enthalpy $H(N, V_{\rm g})$ calculated in the model as a function of $V_{\rm g}$.

and drain, a finite conductance appears. Then a conduction appears outside the yellow parallelogram, inside which the conduction is Coulomb blocked. The parallelogram is called **Coulomb diamond** [*1].

The color plot in Fig. 8.29 shows an example of measured Coulomb diamonds. The sizes of the diamonds are not the same mostly because of the quantum confinement effect explained in the next section. There also should be the variation in the dot size by $V_g$ and the variation in the effective capacitance. Various other effects should affect the sizes of the diamonds and conversely from the size we can know various physical propaties of the dot[2].

We see some line structures outside the diamonds, which come from the quantum confinement and the level descreteness (next section). There are also vague tile-like structures outside the diamond, which is called Coulomb staircase and due to the increase in the number of possible chemical potential levels. In the experiment, we also see that the vertical boundaries have a bit slanted. This is due to the capacitance between the source electrode and the dot. The capacitance mediates some of the electric force lines from the source to the dot.

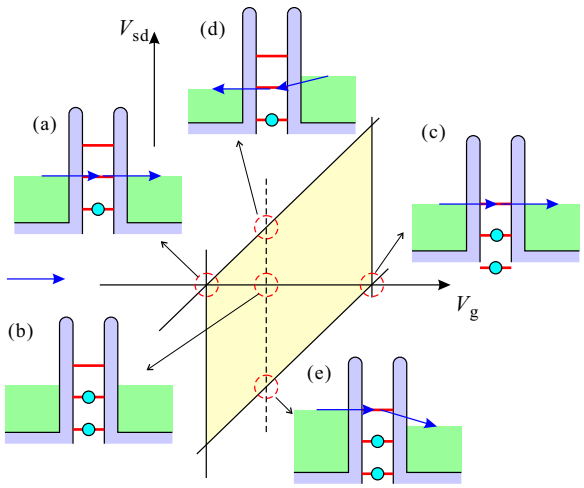



**Fig.** 8.28 Simple description of Coulomb diamonds. In the present case the bias voltage is applied to the source electrode (the drain is grounded). Yellow colored region is a Coulomb diamond (Coulomb blocked region). (a), (c) At zero-bias condiction, electric conduction occurs with tuning the dot chemical potential through the gate voltage. In (b), the system is out of the above resonance condition and the conduction is Coulomb blocked. At finite bias voltages on the source, the conduction appears, in (d) with the use of the chemical potential position that used in (a), vice versa in (e). Finally, the conduction is prohibited in the yellow colored region.

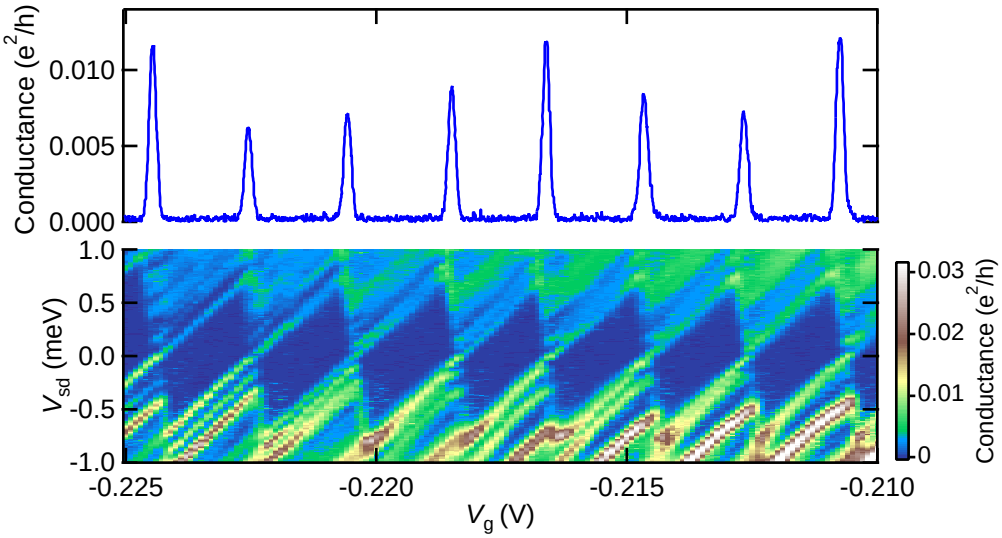



**Fig.** 8.29 Coulomb diamond structure appeared in the transport through a quantum dot made from 2DEG at a heterointerface. The abscissa is the gate voltage $V_g$. The upper panel shows the zero-bias conductance, which shows Coulomb oscillation. In the lower, the conductance is color plotted on the plane of $V_g$-$V_{sd}$. Clear diamond structures are observed. The parallel lines outside the diamonds are from the conduction through excited states in the dot.

[*1] The parallelogram becomes a diamond for symmetric configuration of voltage sources.

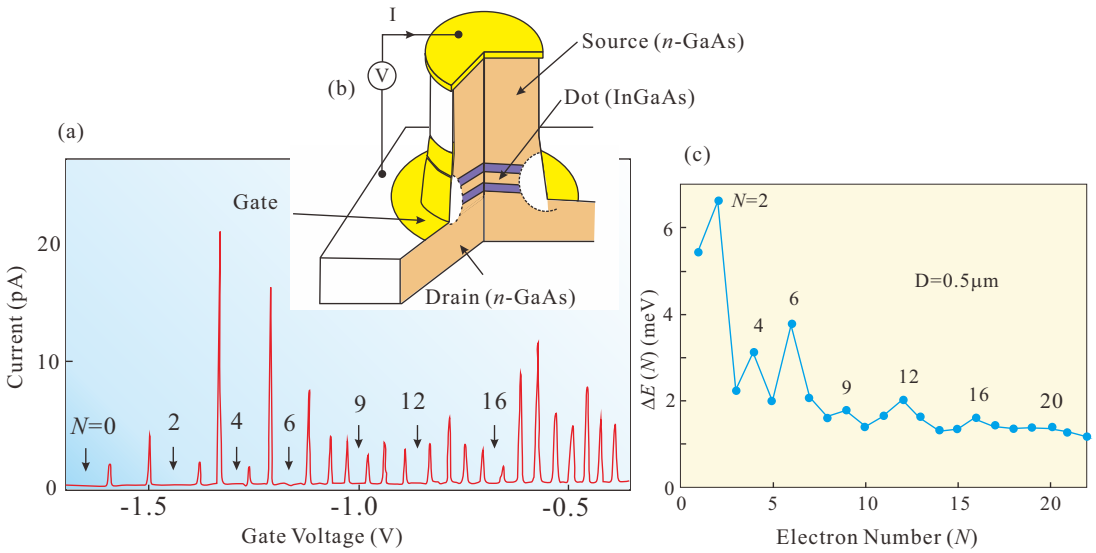



**Fig.** 8.30 Coulomg oscillation in a vertical type quantum dot. (a) Dot current as a function of the gate voltage. There is no Coulomb peak for further negative $V_{\rm g}$ than the region indicated as $N = 0$. The inset illustrates the sample structure. (b) Distances between the Coulomb peaks as a function of the electron number. The data are from [5].

### 8.7.3 Quantum confinement

Next we consider the case we cannot ignore the descreteness of the orbital energy due to quantum confinement. We number the orbital energy levels from the ground state (different numbers are assigned to all the degenerate states). Let the energy of $i$-th state be $\epsilon_i$. We also ignore the terms that do not have relation with $N$. Then enthalpy $H$ is

$$H(N) = \frac{(Ne - C_{\rm g}V_{\rm g})^2}{2C_s} + \epsilon_N. \tag{8.64}$$

The crossing points are obtained as

$$\begin{aligned} \Delta H(N, N+1) = H(N+1) - H(N) = \frac{e}{C_s}\left\{\left(N + \frac{1}{2}\right)e - C_{\rm g}V_{\rm g}\right\} + \Delta\epsilon_N \qquad \Delta\epsilon_N \equiv \epsilon_{N+1} - \epsilon_N \\ V_{{\rm g}X}(N, N+1) = \frac{1}{C_{\rm g}}\left\{\left(N + \frac{1}{2}\right)e + \frac{C_s}{e}\Delta\epsilon_N\right\}, \end{aligned} \tag{8.65}$$

which has a shift from the Coulomb peak position $\Delta\epsilon_N$ as in Eq. (**??**). From the shift we can get the energy spectrum in the quantum dots. This method is called addition energy spectroscopy. Because in the case of degeneracy, $\Delta\epsilon_N = 0$ and from the position of Kramers degeneracy, we can perform quantitative spectroscopy with this as a standard.

Let us have a look on a famous example, in which the researchers realized a two-dimensional harmonic potential. In this experiment, a two-dimensional quantum well was inserted into barrier layers and metallic doped "electrode" layers (source and drain) were placed at the top and the bottom of a cylindrical specimen (vertical type quantum dot). The confinement along vertical direction is strong and we only consider the ground state for this direction. Figure 8.30(a) shows the Coulomb oscillation and there is no peak in the left side (negative $V_{\rm g}$) of a small peak at about $-1.6$ V, which fact indicates that the dot is empty in this region. Then we can assign $N = 0$ to this region and then we can also assign the other number of electrons to blockade regions.

We review two-dimensional harmonic oscillator shortly. The two dimensional coorinate is take to $xy$. The in plane confinement potential $V(x, y)$ and the descrete eigenenergies can be written with the parameter $\omega_0$ representing the strength of the potential as

$$V(x, y) = \frac{m\omega_0^2}{2}(x^2 + y^2), \qquad E_{n_h} = \hbar\omega(n_h + 1) \ \ (n_h = 0, 1, 2, \cdots), \tag{8.66}$$

which have an equidistance. Bound eigenstates of an isotropic potential can be indexed by the quantum number of angular momentum $l$ and the radial quantum number $n_r$. In the above case $n_h = 2n_r + |l|$. The number of possible combinations $(n_r, l)$ is $n_h + 1$, then with spin degeneracy, $E_{n_h}$ has $2(n_h + 1)$ fold degeneracy.

As a simplest analysis of the data in Fig. 8.30(a), the peak intervals are plotted versus the number of electrons $N$ in Fig. 8.30(b). Clear peaks are observed at $N = 2, 6, 12$. This reflects the fact that the bound states in two dimensional harmonic potential take shell structures at $\sum_{j=0}^{n_h} 2(n_h + 1) = (n_h + 1)(n_h + 2)$. In the simplest model, the shift of Coulomb peaks should correspond to $\hbar\omega_0$, hence the peak height in Fig. 8.30(a) should be common. In the experimental data, however, there is a strong tendency that the peak interval decrease with the number of electrons. The tendency is considered to be mainly due to the increase in the effective capacitances. We also see small peak structures at the middle of the clear peaks $N = (n_h + 2)^2$ (4,9). This comes from **Hund's rule**, which tells that the states should be occupied by electrons as to maximize the total spin due to the exchange energy. With quantum dots we can perform experiments knowing the number of electrons and informaiton on potential, quantum dots are sometimes called as "artificial atoms."

Let us see the effect of magnetic field vertical to the 2d-plane. Here we ignore the Zeeman effect. The effect of magnetic field on the orbitals appears in two terms in Hamiltonian. First is the inner product of angular momengum $\boldsymbol{l}$ and the field flux density vector. Second is the confinement into two-dimensional harmonic potential due to the cyclotron motion. The second effect modifies the effective confinement potential as

$$V_{\text{eff}}(x, y) = \frac{m\Omega^2}{2}(x^2 + y^2), \qquad \Omega \equiv \sqrt{\omega_0^2 + (\omega_c/2)^2}, \tag{8.67}$$

where $\omega_c = eB/m$ is the cyclotron frequency for magnetic flux density $B$. Then the energy corresponds to $(n_r, l)$ is

$$E(n_r, l) = \hbar\Omega(2n_r + |l| + 1) + \hbar\omega_c l/2. \tag{8.68}$$

For general finite fields the orbital degeneracy is lifted by the angular momentum. The eigenstates with energies in (8.68) are called **Fock-Darwin states**. The energies in Eq. (8.68) vary with magnetic field as plotted in Fig. 8.31(a). We write

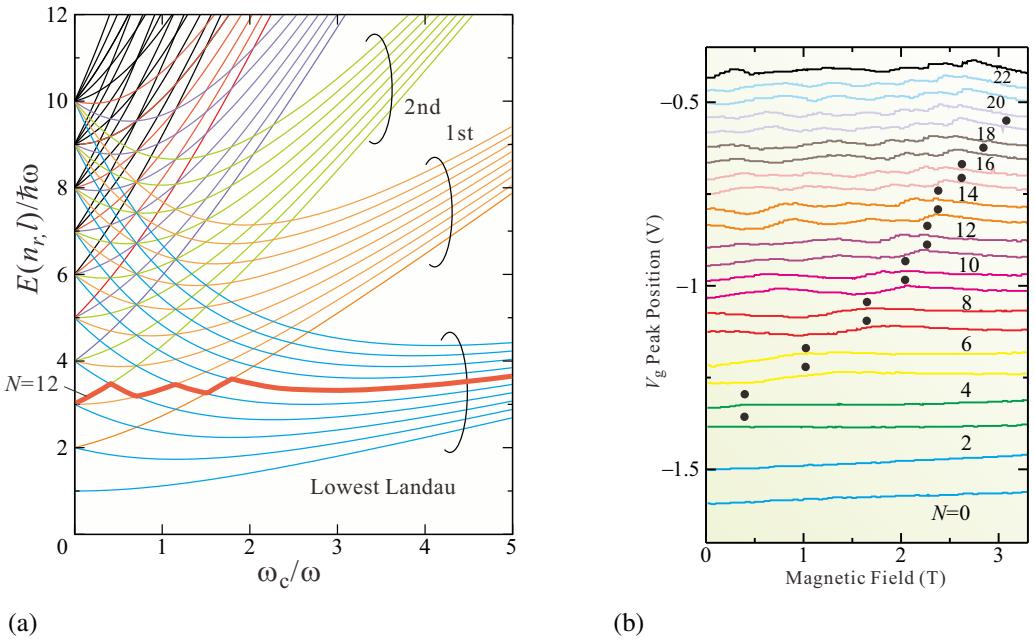



**Fig.** 8.31 (a) Energy levels of Fock-Darwin states (from the ground state to 10th excited state at zero field) as a function of magnetic field. They converge into Landau levels at high fields. Line colors are assigned from corresponding Landau levels. Thick red line is the trace of ground state for electron number $N = 12$. (b) Coulomb peak positions of the quantum dot in Fig. 8.30 versus vertial magnetic field. Black dots are calculated from Eq. (8.69), which represents the position of last crossing point for fixed $N$. The potential parameter $\omega$ is determined from the peaks $N = 3 \sim 6$.

$n_{\rm L} \equiv n_r + (|l| + l)/2$ and take the limit $B \to \infty$, to obtain $E(n_r, l) \to \hbar\omega_c(n_{\rm L} + 1/2)$. That is, they converge into **Landau quantized** levels.

As shown in Fig. 8.31(a), the levels depend on magnetic field with many crossings. The ground state of electrons with a fixed number is given by packing electrons from lower levels. The line of topmost level accommodating electrons should have kinks at such crossings. In Fig. 8.31(a), one of such a line is indicated as a thick red line for $N = 12$. The series of kinks ends up at the field where all the electrons are accomodated into the states corresponding to the lowest Landau level. The last crossing is between the line going to the lowest Landau level for $N$ and the line for $(n_r, l) = (0, 1)$. Because of the spin degeneracy (ignoring Zeeman splitting), the former is given as $(n_r, l) = (0, -\mathrm{int}(N/2))$ ($\mathrm{int}(x)$ is the largest integer smaller than or equal to $x$). This condition is given as

$$2\hbar\Omega + \hbar\omega_c/2 = \hbar\Omega(\mathrm{int}(N/2) + 1) - \hbar\omega_c \mathrm{int}(N/2)/2$$
$$\therefore \quad \left(\frac{\omega_c}{\omega}\right)^2 = \mathrm{int}(N/2) - 2 + \frac{1}{\mathrm{int}(N/2)}. \tag{8.69}$$

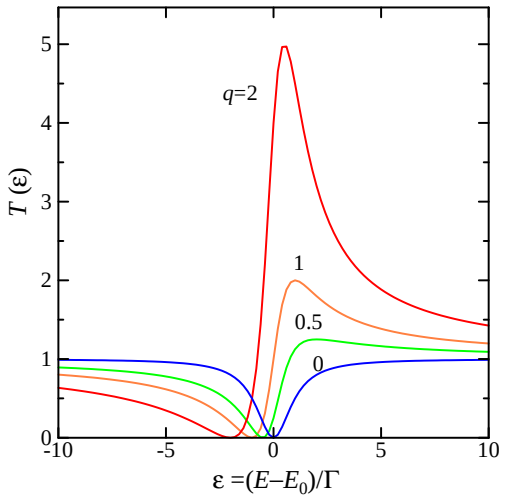
In the first approximation the Coulomb peak distance is constant and ignoring the last term for large $N$, the last crossing points depends on magnetic field prabolically. Actually such behavior is observed in Fig. 8.31(b). If we determine $\omega$ to fit Eq. (8.69) to kinks of $N = 3 \sim 6$, and put dots to the predicted end points of kinks, they agree nicely with the data up to $N = 14$.

## 8.8 Quantum dots and quantum circuits

Quantum dots (QDs) can be connected with quantum wires to form quantum circuits. A QD affects the circuit conductance though the transmission probability and the phase shift as characteristics of resonant scattering. Here the effect of single electron charging is only on the positions of chemical potentials for resonant scatterings. Hence in the simplest approximation, QDs are treated simple resonators and we do not consider the single electron effect explicitly.

### 8.8.1 Quantum dot and scattering experiment

Transport in mesoscopic systems can be viewed as scattering experiments in solids. We can see that clearly in quantum circuits with QDs.




For example, in the scattering of electrons with an atom, the **Fano resonance** occurs as the interference between the incident wave with continuous energies and the wave scattered by discrete atomic levels. And the same effect is observe in circuits with QDs. In the Fano effect, the scattered wave gets rapid phase shift by $\pi$ over the resonance position and the interference results in distorted lineshape in the resonance. There the energy dependence of transmission coefficient is given as

$$T(\tilde{\epsilon}) \propto \frac{(\tilde{\epsilon} + q)^2}{\tilde{\epsilon}^2 + 1}, \quad \tilde{\epsilon} \equiv \frac{E - E_0}{\Gamma}. \tag{8.70}$$

Here $q$ is called Fano parameter, which determines the lineshape as in the left figure. The larger absolute value in $q$ results in the larger asymmetry and $q = 0$ gives a symmetric dip (anti-resonance).

Here we do not go into the derivation of (8.70)[2]. Rather we modelize the circuit with S-matrices and obtain the lineshape numerically. One-dimensional band of quantum wires are assigned to "continuum states" and quantum confined

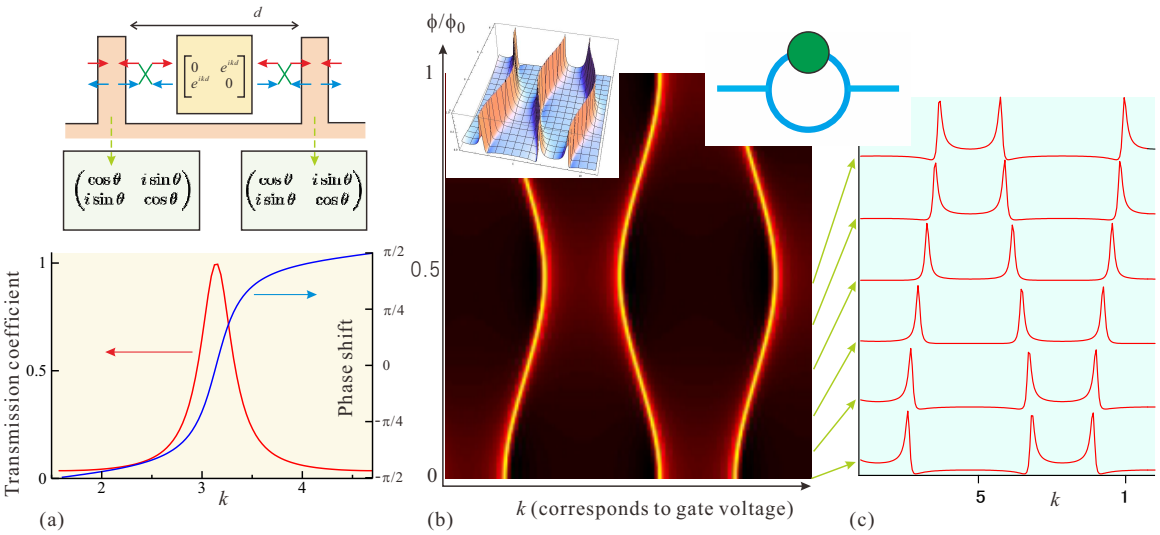



**Fig.** 8.32 (a) Upper: Double barrier model. Lower: Transmission coefficient (red) and phase shift (blue) of the model in the upper panel. $k$ corresponds to gate voltage. The reflaction coefficient of a barrier is 0.7. (b) Color plot of conductance of the system shown in the right upper inset (AB ring+QD) versus plane of $k$ (gate voltage) and magnetic flux $\phi$ piercing the ring. The conductance is higher for black $\rightarrow$ red $\rightarrow$ yellow. In the AB ring model in Fig.8.19(a), $\boldsymbol{S}_{\rm w}$ is replaced with the S-matrix obtained in (a), and a finite transmission coefficient is introduced into $\boldsymbol{S}_{\rm AB}$. Reflection at the dot barrier is 0.7 and that in the reference arm is 0.82. (c) Transmission of (b) is plotted versus $k$ for $\phi/\phi_0$ =0, 0.01, 0.19, 0.29, 0.38, 0.48 from down to up.

discrete states in a QD are assigned to "discrete states." To have interference between incident wave and scattered wave, the incident wave is devided into two, one of which goes through a QD and the ohter directly goes to the outlet. A QD is formed as a one-dimensional double barrier structure. The model is described by S-matrices as in Fig. 8.32(a). That is, the barrier and the dot S-matrices are

$$S_{\rm b} = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix}, \quad S_{\rm d} = \begin{pmatrix} 0 & e^{ikd} \\ e^{ikd} & 0 \end{pmatrix}. \tag{8.71}$$

Here $k$ is the wavenumber representing kinetic energy, which corresponds to a gate voltage. In this model, the transmission coefficient and the phase shift are calculated from the composite S-matrix as shown in Fig. 8.32(a), where $\pi$ phase shift at the resonance peak is clearly observed. This is common for resonance. Resonance is a response of system, in which a pole exists in the region $\mathrm{Re}(z) < 0$ on the complex $z$-plane. The angle from the pole to a point on the real axis changes from $-\pi$ to 0 with the movement of the point from $-\infty$ to $+\infty$.

As in Sec. 8.5.4, an S-matrix for two junctions with three channels is written as

$$S_t = \begin{pmatrix} 0 & -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & -1/2 \\ -1/\sqrt{2} & -1/2 & 1/2 \end{pmatrix}. \tag{8.72}$$

Also for the AB phase, we insert an S-matrix

$$S_{\rm AB} = \begin{pmatrix} 0 & e^{i\theta_{\rm AB}} \\ e^{-i\theta_{\rm AB}} & 0 \end{pmatrix}, \qquad \theta \equiv 2\pi\frac{\phi}{\phi_0} = \frac{e}{\hbar}\phi \ \ (\phi \text{ is flux through the ring.}) \tag{8.73}$$

to one of the parallel paths. And the QD represented in (8.71) is inserted into the other path. Though thus obtained analytical form of transmission coefficient is complicated, numerical calculations shows clear Fano effect as in Fig. 8.32(b), (c). The direction of the lineshape distortion (parameter $q$ in (8.70)) changes with the period $\phi_0$ as shown in (c). This is natural consequence of the interference and evidences that the distortion comes from the rapid $\pi$ change in the phase shift appeared in Fig. 8.32(a).

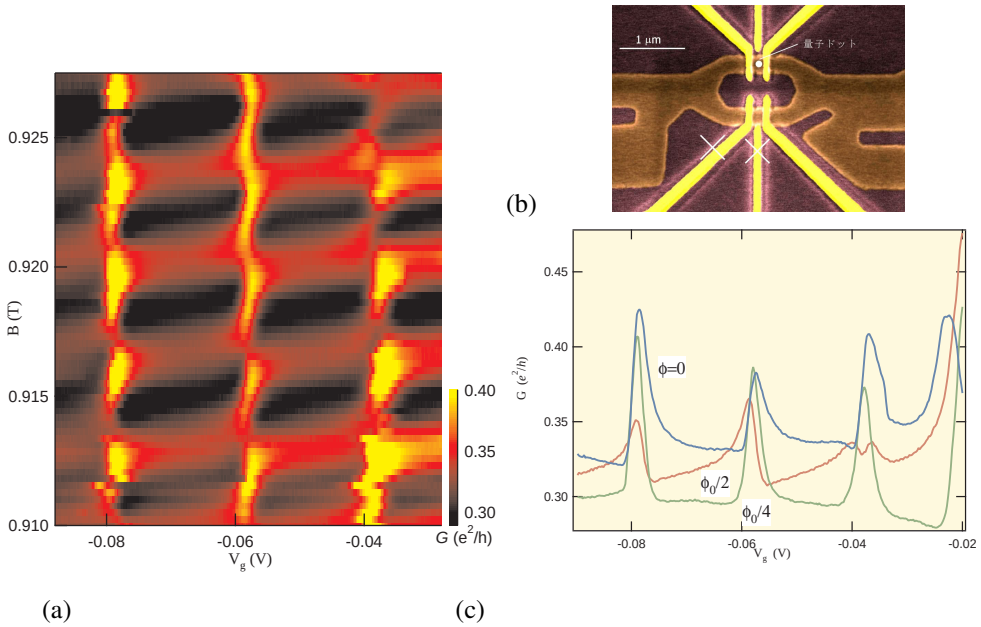



**Fig.** 8.33 (a) Color plot of conductance of an AB ring with a QD on one of the paths versus gate voltage and magnetic field. (b) Scanning electron micrograph of the sample. The gates with $\times$ mark are not used. (c) Coulomb oscillations at three representative magnetic fields[1].

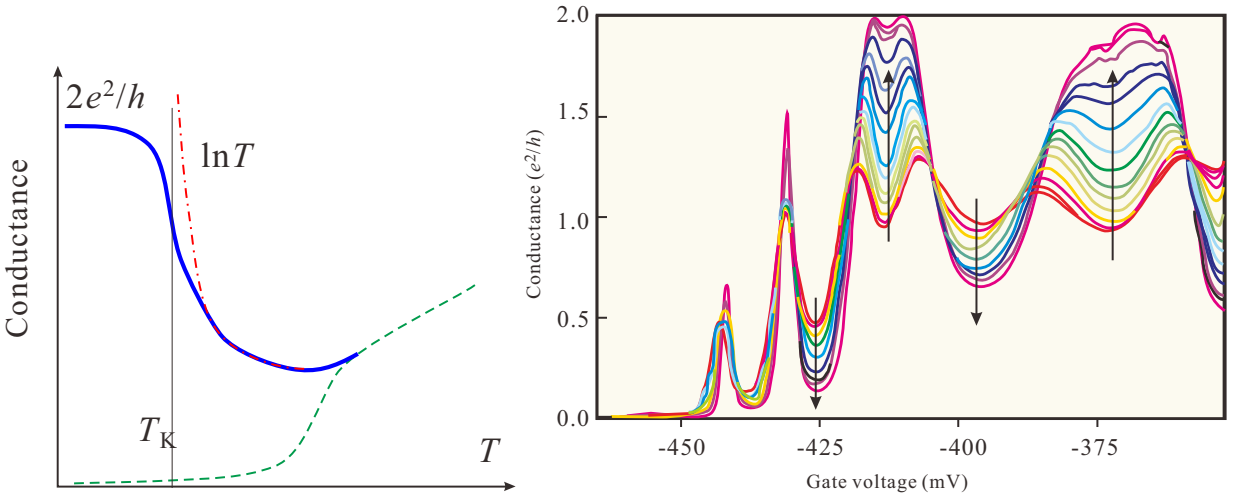
Figure 8.33 shows results of an experiment. The Fano lineshape and Fano parameter vary aginst the flux piercing the AB ring as expected.

### 8.8.2 Qauntum dot and the Kondo effect

The Kondo effect in QDs originates from the indirectinteraction between electrons in electrodes via localized states. We do not go into the details due to time limitation. In Appendix 8C, very short summary is given. At low temperatures most of the freedoms die out and conductors are like empty cavities (electron "waveguides" are also cavities just like microwave wavefuides). An exceptional case comes from the existence of energy-degenerated freedoms just at the Fermi level. Fermi spheres themselves are such exceptions but if there exists another degenerated freedom exists and the freedom has quantum entanglement with electrons at the Fermi level, the **Kondo effect** appears. The Kramers degeneracy due to time-reversal symmetry, i.e. spin degeneracy is an easiest example of such degenerated freedom. Hence QDs with odd number of electrons are convenient for the experiments because the topmost level is occupied by a single electron and has spin 1/2.

The Kondo effect first appeared as increase of resistance in diluted magnetic alloys with decreasing temperature. Jun Kondo gave theoretical solution to this problem and simultaneously found the divergence in the second order perturbation. This **Kondo problem** became a big problem of physics beyond the frame of solid state physics. For the problem, Anderson impurity model was proposed, **renormalization group theory** was developed. The renormalization group theory was applied to quark confinement problem in particle physics and led to the concept of asymptotic freedom and the establishment of quantum chromodynamics.

As shown in App. 8C, in very short, the Kondo effect in QDs is anomalous enhancement of tunneling probability from Hamiltonian $H_{\rm T}$ by many body resonance between degenerate freedom and the Fermi surface. In the case of QDs, the process expressed by $H_T$ is the transmission of electrons through the QDs. That is, enhancement of $H_T$ means enhancement of conductance. If we consider the anaoly with double barrier resoance, the Kondo many body resonance anomalously enhances the conductance and even when the original conductance is very small due to the




(a) (b)

**Fig.** 8.34 (a) Illustration of electric conductance through a quantm dot as a function of temperature with the emergence of the Kondo effect. Solid blue line is for odd number of electrons in the QD. In this case the Kondo effect emerges and at high temperatures the conductance is enhanced in proportion to $\ln T$ as the temperature decreases, whereas below $T_{\rm K}$ the enhancement is saturated to reach the unitary limit $2e^2/h$. Green broken line is for even number of electrons and the conductance goes to zero with the Coulomb blockade. (b) The Kondo effect appeared in the conduction experiment of a QD. The parameter is temperature and the arrows indicate the direction of lowering the temperature, with which the conductance increases in the valleys with odd nubmer of electrons and decreases in those with even number of electrons[6].

Coulomb blockade, the final transmission probability should be 1 (unitary). In this case, the conductance is, from the Landauer formula, the universal value $2e^2/h$.

A characteristic feature of the Kondo effect is that it is always in resonance with the Fermi surface. Therefore roughly speaking in the Coulomb valleys with odd number of electrons the conductance is $2e^2/h$ and in those with even number of electrons the conductance is zero. In Fig. 8.34, we show conceptual behavior and actual observation of QD conductance around the temperature characteristic of the Kondo effect (Kondo temperature, $T_{\rm K}$).

## Appendix 8C: The Kondo effect

We consider a QD with the impurity Anderson model, in which only one electrons exists in the dot ($n = 1$) as the ground state. The dot has spin 1/2, can be viewed as a kind of magnetic impurity. We write the energy required to add an electron as $\Delta E^+$, that required to extract one as $\Delta E^-$. Then

$$\Delta E^+ = \mu_2 - \mu = \epsilon_0 + U - \mu, \quad \Delta E^- = \mu - \mu_1 = \mu - \epsilon_0. \tag{8C.1}$$

These energies give the state-allowance-times $h/\Delta E^\pm$ from the uncertainty relation for the excited states. There should be, then, second order tunneling processes with $H_T$ by utilizing these excited states as the intermediate states. The probabilities of such processes are

$$\frac{-\gamma_L^*\gamma_R}{\Delta E^-}, \quad \frac{\gamma_L^*\gamma_R}{\Delta E^+}. \tag{8C.2}$$

Such tunnel processes of higher order is called co-tunneling. The Kondo effect can also be regarded as a phenomenon in which the tunnel probability amplitude due to co-tunneling increases anomalously.

First, the Hamiltonian

$$H = H_{\rm leads} + H_{\rm dot} + H_{\rm T} \tag{8C.3}$$

is unitary-transformed as

$$\begin{cases} c_{k\sigma} = (\gamma_L^* c_{L,k\sigma} + \gamma_R^* c_{R,k\sigma})/\gamma \\ \bar{c}_{k\sigma} = (-\gamma_R c_{L,k\sigma} + \gamma_L c_{R,k\sigma})/\gamma \end{cases}, \quad \gamma^2 \equiv \gamma_L^2 + \gamma_R^2 \tag{8C.4}$$

Then the tunnel Hamiltonian is transformed as

$$\begin{aligned} H_T &= \sum_{k,\sigma} [(\gamma_L c_{L,k\sigma}^\dagger + \gamma_R c_{R,k\sigma}^\dagger) d_\sigma + \text{h.c.}] \\ &= \sum_{k,\sigma} [\gamma c_{k\sigma}^\dagger d_\sigma + \text{h.c.}], \end{aligned} \tag{8C.5}$$

in which we can ignore $\bar{c}_{k\sigma}$ because it has nothing to do with the coupling to the dot. This transformation renormalizes the two electrodes model of QD into "a QD and a system with a Fermi surface" model. It formally equalizes a QD for transport experiment with electrodes to a magnetic impurity in a metal [*2].

The transformed Anderson impurity model Hamiltonian is written as

$$H = \sum_{k\sigma} \epsilon_k c_{k\sigma}^\dagger c_{k\sigma} + \sum_\sigma \epsilon_d d_\sigma^\dagger d_\sigma + \sum_{k\sigma} (\gamma c_{k\sigma}^\dagger d_\sigma + \text{h.c.}) + U d_\uparrow^\dagger d_\uparrow d_\downarrow^\dagger d_\downarrow. \tag{8C.6}$$

The condition for having single electron in the ground state of the dot is

$$\epsilon_d < E_\text{F} < \epsilon_d + U. \tag{8C.7}$$

Under the condition, we regard the interaction term (the third term with $V_{kd}$) of the conduction electron ($s$-electron) in the electrode and the dot electron ($d$-electron) as a perturbation. The first order of perturbation does not exist because it changes the number of $d$-electron, and the leading order is second. This means we need to consider co-tunneling process as of the leading order.

There are following four perturbation processes on the state in which the $d$-electron has up-spin $\uparrow$. The contraint is that only $\downarrow$-electron is allowed to enter the dot by Pauli principle. We write the unperturbed state as $\psi_\uparrow$.

1) $k\downarrow \to d\downarrow \to k'\downarrow$

2) $k\downarrow \to d\downarrow$, $d\uparrow \to k'\uparrow$ (down-spin electron goes into the dot then up-spin electron goes out)

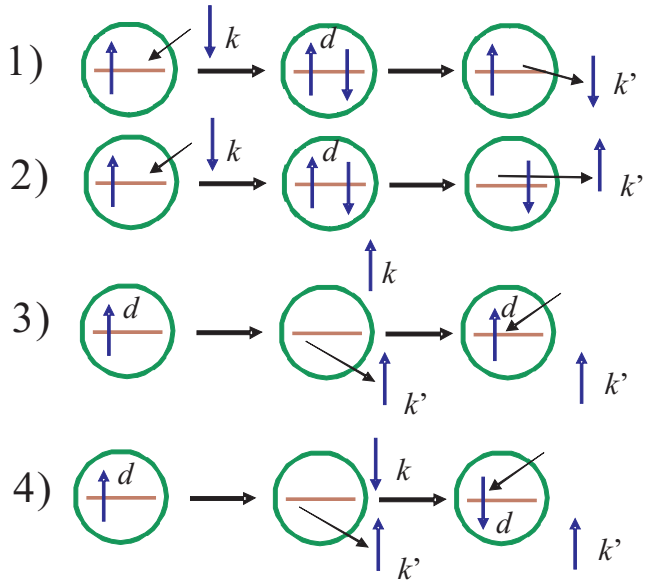



**Fig.** 8C.1 2nd order possible tunneling processes for the initial states of an up-spin electron inside the dot These 1)∼4) correspond to 1)∼4) in Eq. (8C.8), (8C.9) respectively.

[*2] There is a difference in the physical meaning. In the case of magnetic impurities, $c_{k'\sigma}^\dagger c_{k\sigma}$ means impurity scattering. In the case of QDs, on the other hand, it represents transmission and reflection via co-tunneling. Another difference is that in the case of impurities, $k$ should be three dimensional vectors. This can be, however, transformed to one-dimensional problem with partial wave expansion of scattered wave, and mathematical equivalence is kept. This method is used to find the exact solution based on the Bethe ansatz.

3) $d\uparrow \to k'\uparrow$, $k\uparrow \to d\uparrow$ (up-spin electron goes out then up-spin electron goes into the dot)

4) $d\uparrow \to k'\uparrow$, $k\downarrow \to d\downarrow$ (up-spin electron goes out then down-spin electron goes into the dot)

Effective Hamiltonians for these processes are

$$1) \to -\frac{\gamma^2}{\Delta E^+} c^\dagger_{k'\downarrow} d_\downarrow d^\dagger_\downarrow c_{k\downarrow}, \quad 2) \to -\frac{\gamma^2}{\Delta E^+} c^\dagger_{k'\uparrow} d_\uparrow d^\dagger_\downarrow c_{k\downarrow}, \tag{8C.8}$$

$$3) \to \frac{\gamma^2}{\Delta E^-} d^\dagger_\uparrow c_{k\uparrow} c^\dagger_{k'\uparrow} d^\dagger_\uparrow, \quad 4) \to \frac{\gamma^2}{\Delta E^-} d^\dagger_\downarrow c_{k\downarrow} c^\dagger_{k'\uparrow} d_\uparrow. \tag{8C.9}$$

Just same as above, four perturbation processes exist for the case that $\psi_\downarrow$ is the initial state. The effective Hamiltonians for these processes are obtained by the replacement $\uparrow\downarrow \to \downarrow\uparrow$. These terms are summed up to be

$$\begin{aligned} &\sum_{k\sigma} \frac{\gamma^2}{\Delta E^-} d^\dagger_\sigma d_\sigma + \sum_{kk'\sigma} \frac{\gamma^2}{\Delta E^+} c^\dagger_{k'\sigma} c_{k\sigma} \\ &+ \sum_{kk'} \gamma^2 \left( \frac{1}{\Delta E^+} + \frac{1}{\Delta E^-} \right) (c^\dagger_{k'\uparrow} c_{k\uparrow} d^\dagger_\uparrow d_\uparrow + c^\dagger_{k'\downarrow} c_{k\downarrow} d^\dagger_\downarrow d_\downarrow + c^\dagger_{k'\uparrow} c_{k\downarrow} d^\dagger_\downarrow d_\uparrow + c^\dagger_{k'\downarrow} c_{k\uparrow} d^\dagger_\uparrow d_\downarrow). \end{aligned} \tag{8C.10}$$

The first term represents process 3) for the case of $k = k'$. Because $k$ is outside the Fermi surface, at low temperature under Fermi degeneracy condition, we assume $c_k c^\dagger_k = 1$, $c^\dagger_k c_k = 0$. The second term is for process 1). To obtain this term we use the fact that from $d_\downarrow \psi_\uparrow = 0$, we can write $d_\downarrow d^\dagger_\downarrow = 1$, $d^\dagger_\downarrow d_\downarrow = 0$. The residual part of process 3) and those of 2) and 4) are expressed in the third term.

Here we transform the above to

$$c^\dagger_{k'\uparrow} c_{k\uparrow} d^\dagger_\uparrow d_\uparrow + c^\dagger_{k'\downarrow} c_{k\downarrow} d^\dagger_\downarrow d_\downarrow = \frac{1}{2}(c^\dagger_{k'\uparrow} c_{k\uparrow} - c^\dagger_{k'\downarrow} c_{k\downarrow})(d^\dagger_\uparrow d_\uparrow - d^\dagger_\downarrow d_\downarrow) + \frac{1}{2}(c^\dagger_{k'\uparrow} c_{k\uparrow} + c^\dagger_{k'\downarrow} c_{k\downarrow})(d^\dagger_\uparrow d_\uparrow + d^\dagger_\downarrow d_\downarrow).$$

Because the spin operator of the dot $\hat{S}$ is expressed as

$$\hat{S}_z = \frac{1}{2}(d^\dagger_\uparrow d_\uparrow - d^\dagger_\downarrow d_\downarrow), \quad \hat{S}_+ = d^\dagger_\uparrow d_\downarrow, \quad \hat{S}_- = d^\dagger_\downarrow d_\uparrow,$$

the summation of the second and the third term in (8C.10) is rewritten to the summation of the following two Hamiltonians ($H_d$,$H_{sd}$):

$$H_d = \sum_{kk'\sigma} \gamma^2 \left[ \frac{1}{\Delta E^+} - \frac{1}{2} \left( \frac{1}{\Delta E^+} + \frac{1}{\Delta E^-} \right) \right] c^\dagger_{k'\sigma} c_{k\sigma}, \tag{8C.11}$$

$$H_{sd} = \sum_{kk'} \gamma^2 \left[ \frac{1}{\Delta E^+} + \frac{1}{\Delta E^-} \right] \left[ \hat{S}_+ c^\dagger_{k'\downarrow} c_{k\uparrow} + \hat{S}_- c^\dagger_{k'\uparrow} c_{k\downarrow} + \hat{S}_z (c^\dagger_{k'\uparrow} c_{k\uparrow} - c^\dagger_{k'\downarrow} c_{k\downarrow}) \right]. \tag{8C.12}$$

Let us define $J$ as

$$J = \gamma^2 \left( \frac{1}{\Delta E^+} + \frac{1}{\Delta E^-} \right), \tag{8C.13}$$

then

$$H_d = \sum_{kk'} \left( -\frac{J}{2} \right) c^\dagger_{k'\sigma} c_{k\sigma} \tag{8C.14}$$

is ordinary potential scattering, which does not depend on spin. On the other hand,

$$\begin{aligned} H_{sd} &= J \sum_{kk'} \left[ \hat{S}_+ c^\dagger_{k'\downarrow} c_{k\uparrow} + \hat{S}_- c^\dagger_{k'\uparrow} c_{k\downarrow} + \hat{S}_z (c^\dagger_{k'\uparrow} c_{k\uparrow} - c^\dagger_{k'\downarrow} c_{k\downarrow}) \right] \\ &= J \sum_j \left[ (\hat{S}_x + i\hat{S}_y)(\hat{s}_{xj} - i\hat{s}_{yj}) + (\hat{S}_x - i\hat{S}_y)(\hat{s}_{xj} + i\hat{s}_{yj}) + 2\hat{s}_{zj}\hat{S}_z \right] \\ &= 2J \sum_j \hat{\boldsymbol{s}}_j \cdot \hat{\boldsymbol{S}} \end{aligned} \tag{8C.15}$$

is expressing the exchange interaction between spin of conduction electrons $\boldsymbol{s}_j$ and spin on the dot. This is often called $sd$-Hamiltonian, which originally expresses interaction between electron spin ($s$) and localized spin (in many cases, in

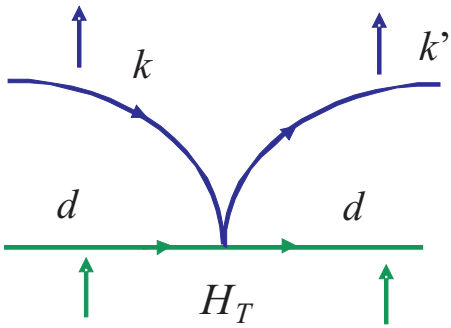



**Fig.** 8C.2 Diagram representing electron scattering $k \to k'$ by a dot in the first order of $H_T$. The time flows from the left to the right. $d$ represents the dot and the up/down arrows indicate spins.

$d$-orbital, then $d$-spin) in diluted magnetic impurity system. Now with unitary transformation in (8C.4), we can also apply the $sd$-Hamiltonian to QD-electrode systems.

Then as transmission Hamiltonian $H_T$ we ignore potential scattering $H_d$ and take only $H_{sd}$. Then with adding the electorons in electrodes, the effective Hamiltonian

$$H_{\text{eff}} = \sum_{k\sigma} \epsilon_k c^\dagger_{k\sigma} c_{k\sigma} + J \sum_{kk'} \left[ \hat{S}_+ c^\dagger_{k'\downarrow} c_{k\uparrow} + \hat{S}_- c^\dagger_{k'\uparrow} c_{k\downarrow} + \hat{S}_z (c^\dagger_{k'\uparrow} c_{k\uparrow} - c^\dagger_{k'\downarrow} c_{k\downarrow}) \right]. \tag{8C.16}$$

is obtaiend (Schrieffer-Wolff transformation).

Kondo calculated the scattering amplitude by the effective Hamiltonian (8C.16) to the second order in Born approximation. That is, he treated $J$ in (8C.16) as the parameter of perturbation and calculated up to the quadratic term of $J$ (the forth term of $\gamma$). The operator of transition between the left and right electrodes is given as

$$\hat{T} = H_T + H_T \frac{1}{\epsilon - H_0 + i\delta} H_T + \cdots . \tag{8C.17}$$

The tunnel probability of $L \to R$ is formally written as

$$\Gamma_{L\to R} = 2 \sum_{k,k'} \frac{2\pi}{\hbar} \left| \langle Rk' | \hat{T} | Lk \rangle \right|^2 \delta(\epsilon_{Rk'} - \epsilon_{Lk}) f(\epsilon_{Lk} - \mu_L)[1 - f(\epsilon_{Rk'} - \mu_R)]. \tag{8C.18}$$

Let us treat the scattering $|k\uparrow\rangle \to |k'\uparrow\rangle$. Perturbation to the first order of $J$ is expressed in the diagram shown in Fig. 8C.2 and calculated as

$$\langle d\uparrow; k'\uparrow | \hat{T}^{(1)} | d\uparrow; k\uparrow \rangle = J/2. \tag{8C.19}$$

The conduction process $H_T$ requires two consecutive tunnelings and this corresponds to the second order of $\gamma$ ($J$ is thus proportional to $\gamma^2$), and co-tunneling process in (8C.2).

There are three types of processes in the second order of $J$ $\langle d\uparrow; k'\uparrow | \hat{T}^{(2)} | d\uparrow; k\uparrow \rangle$ as shown in Fig. 8C.3 and Fig. 8C.4. The first and second processes are not associated with spin flip and they are distinguished as electron process (Fig. 8C.3(a)) and electron-hole pair process (Fig. 8C.3(b)) for the intermediate propagation process [*3] . The contribution of these two terms is calculated as

$$\begin{aligned}
&\sum_q \left(\frac{J}{2}\right)^2 \frac{1}{\epsilon - \epsilon_q + i\delta}[1 - f(\epsilon_q)] + \sum_q \left(\frac{J}{2}\right)^2 \frac{-1}{\epsilon - (2\epsilon - \epsilon_q) + i\delta} f(\epsilon_q) \\
&= \sum_q \left(\frac{J}{2}\right)^2 \frac{1}{\epsilon - \epsilon_q + i\delta} \\
&= \left(\frac{J}{2}\right)^2 \int_{-D}^{D} d\epsilon' \nu \frac{1}{\epsilon - \epsilon' + i\delta} \qquad \nu : \text{Density of states} \\
&= \left(\frac{J}{2}\right)^2 \nu \left[ \ln \left| \frac{D + \epsilon}{D - \epsilon} \right| - i\pi \right].
\end{aligned} \tag{8C.20}$$

[*3] Here "hole" state refers to Fermi liquid lacking single electron. This is largely different from the "hole" state defined as the state created by extracting an electron from valence band (Sec. 3.1.2).

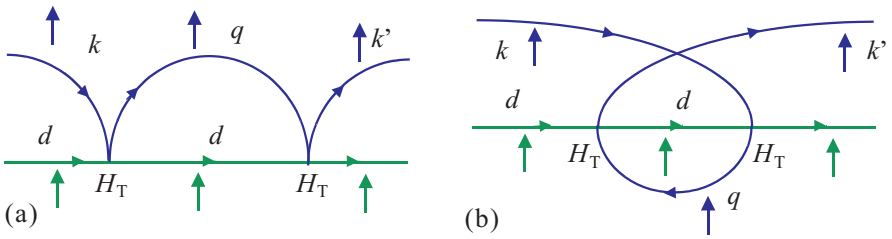



**Fig.** 8C.3 Non spin-flip processes for 2nd order of $H_T$. (a) Process with electron excitation as the intermediate state. (b) In the intermedate state of this process, an electron-hole pair propagates. The hole is annihilated by recombination with an electron in the electrode.

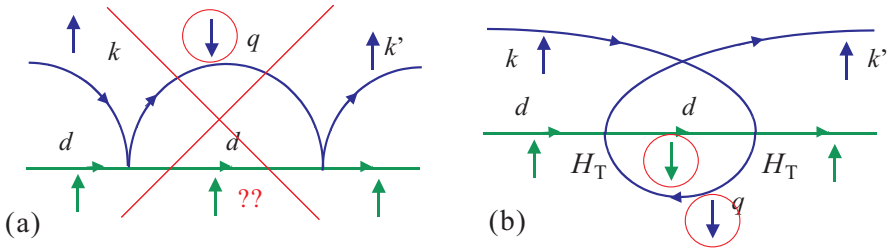



**Fig.** 8C.4 Processes with spin flipped intermediate stats in the second order of $H_T$. (a) Process with electron excitation as the intermediate state. This process is absent due to the angular momentum conservation. (b) In the intermedate state of this process, an electron-hole pair propagates.

The result does not show any anomaly as a consequence of electron-hole symmetry. Here as for the electronic states in metal, we have adopted a rough (or abstract) approximation that a band spread over $[-D, D]$ on energy with a uniform density of states $\nu$. Such a "toy" model is often good to see the essence of phenomenon.

In the processes shown in Fig. 8C.4, spin flips occur in the intermediate states. However in (a), the dot spin should be 3/2 for the conservation of angular momentum, and this process is prohibited. In (b), an electron-hole pair propagates in the intermediate state and the contribution is calculated as

$$\sum_q J^2 \frac{1}{\epsilon - \epsilon_q + i\delta} f(\epsilon_q) = J^2 \nu \int_{-D}^{D} \frac{1}{\epsilon - \epsilon' + i\delta} f(\epsilon') d\epsilon$$
$$\approx \begin{cases} -J^2 \nu \ln |\epsilon|/D & |\epsilon| \gg k_{\rm B}T, \\ -J^2 \nu \ln k_{\rm B}T/D & |\epsilon| \ll k_{\rm B}T. \end{cases} \tag{8C.21}$$

This term diverges logarithmically with temperature lowering or smaller $\epsilon$. This is the anomalous term found by Kondo. And various phenomena originate from this anomaly are called the **Kondo effect**.

Let us consider the origin of this term. In the case of non-spin-flip processes, the anomalous terms cancel each other due to the electron-hole symmetry. That is, if we look electrons or holes separately the anomaly exists regardless of spin-flip and the origin is the existence of Fermi surface, which represents huge asymmetry. At absolute zero, states inside a Fermi sphere are fully occupied while those above the Fermi surface are completely empty with almost infinite degeneracy. In the processes without spin flip, the electron-hole symmetry perfectly cancels this huge asymmetry. On the other hand in the processes with spin flip, conservation of spin angular momentum prohibits electron propagation process in the intermediate state [*4], and the asymmetry at the Fermi surface appears as the anomaly.

Because the perturbation leads to the divergence, the perturbative treatment itself is in failure for the condiitons close

[*4] Mathematically this comes from non-commutativity of $\hat{S}_+$ and $\hat{S}_-$ (the commutation relation yields $\hat{S}_z$), hence one can say that this is due to a quantum mechanical effect.

to the divergence. Treatment of this problem thus requires various methods other than simple perturbation. In order for handling this **Kondo problem**, a number of methods including renormalization group, have been developed. Here we do not go into the detail of such methods. Instead we have a look on the results at lowest temperatures obtained from years of research.

In $sd$ Hamiltonian (8C.15), there is antiferromagnetic interaction between the dot spin and the conduction electron spin. That is, the dot spin attracts electrons with anti-parallel spins and repels those with parallel spins through **exchange interaction** arises from co-tunneling. As a result, seen from a distance, a cloud of spin polarization of conduction electrons appears to cling to the dot spin. The Kondo problem indicates the anomalous enhancement of the above effect. The cloud like state of spin polarization created as above is called **Kondo cloud**. On the other hand, the spin polarization surrounding the dot spin reaches complete screening of the dot spin, further polarization stops. This is called **unitary limit**.

This phenomenon can be viewed as **many-body resoance**. A representative of single-body resonance (resonace that comes from potential) is the resonant tunneling through double barrier structures. In the double barrier structure, no matter how high the barrier height is and how small the tunnel probability is, where the energy of the incident wave is in resonance, the reflected and transmitted waves are infinite sum up of coherent reflection and transmission by the two barriers. And finally the reflections cancel each other out, the total transmittance is 1. This resonance energies are close to the bound state energies of an imaginary quantum well composed by making the barrier thicknesses infinite. When the barrier thickneeses are finite, resonance makes average staying time anomalously long and the modes are called quasi-bound states. Even if an electron enters the quasi-bound state, it eventually leaks to the outside, so it is in a resonance state with the free electrons in the electrodes. When the Fermi level hits resonance, the transmission probability reaches a peak value. An example is shown in Fig. 8C.5.

The Kondo effect has many common features with double barrier phenomenon. While single-body resonance is based on infinite number of reflections, the Kondo resonance occurs as a result of infinite degeneracy at Fermi surface. In potential resonance, orbital effect results in non-uniform probability distribution. In the case of the Kondo effect, the force works among spins and no charge inhomogeneity appears. Instead, localized spin polarization occurs as Kondo cloud. The biggest difference is that the Kondo cloud is always in resonance with Fermi surface.

Electrons stay in quasi-bound states for finite times. Let $\tau_a$ be the average staying time and the resonance has lifetime broeadning $h/\tau_a(=\hbar\Gamma$, $\Gamma$ is tunneling frequency). Does the Kondo cloud have width? If it does, how large is that? When the thermal broadening of Fermi surface is larger than the resonance width, the temperature dependence of the contribution of resonance to conduction should be weak. On the contrary, when the thermal width is narrower than the resonance width, the temperature dependence also disappears. The energy scale representing resonance width is, in temperature unit, called **Kondo temperature**, for which symbol $T_{\rm K}$ is often used.

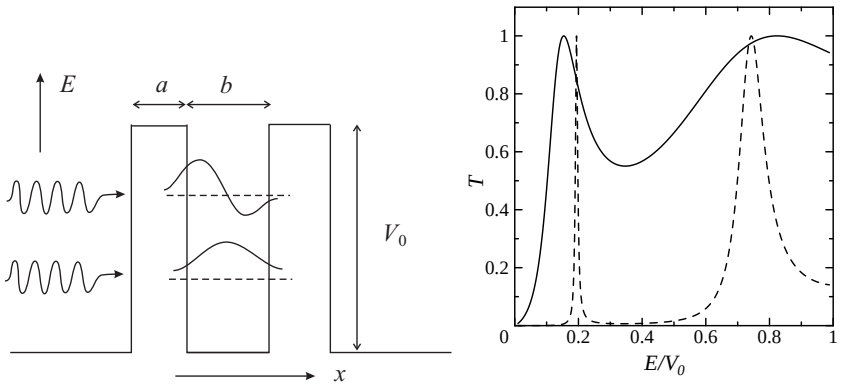



**Fig.** 8C.5 Left: Schematics of double barrier potential. Vertical axis is energy, horizontal axis is space coordinate. Broken lines indicate the positions of quasi-bound state energies, i.e. positions of resonant tunneling. Real wavelengths of incident waves are much longer than the illustration. Right: Example of transmission probability under condition of $k_0 \equiv \sqrt{2mV_0/\hbar}$ and $k_0 b$=5.0. Solid line is for $k_0 a$=0.5 and broken line is for $k_0 a$=2.0.

For correct estimation of Kondo temperature, perturbation in (8C.21) is not enough and the effect of higher order terms should be taken into account in some way. Though, roughly speaking, $T_{\mathrm{K}}$ can be estimated as the temperature-dependent term in (8C.21) is comparable with $J$ at high temperature side. That is, from

$$-J^2\nu \ln k_{\mathrm{B}}T_{\mathrm{K}}/D \sim J,$$

$T_{\mathrm{K}}$ is given as

$$k_{\mathrm{B}}T_{\mathrm{K}} \sim De^{-J\nu}.$$

In the above rough estimation, with increasing the anti-ferromagnetic coupling strength $J$, $T_{\mathrm{K}}$ dcreases exponentially. Larger $J$ corresponds to higher barriers in double barrier, narrower lifetime width with decreasing tunneling probability. The same for the density of states $\nu$. On the other hand, widening of band $D$ makes resonance wider, $T_{\mathrm{K}}$ higher. Actually in the present simple model, $J$ and $D$ are not independent. But with ignoring that, widening of the resonance by increasing $D$ is interpreted as the increase of contribution from deeper inside the Fermi sphere.

## References


[1] K. Kobayashi *et al.*, Phys. Rev. Lett. **88**, 256806 (2002); Phys. Rev. B **68**, 235304 (2003).

[2] 勝本信吾,「半導体量子輸送物性」 (培風館，2014).

[3] 近藤 淳,「金属電子論」(裳華房，1983)，芳田圭,「磁性」(岩波書店，1991)

[4] Yu. V. Nazarov and Y. M. Blanter, “Quantum Transport” (Cambridge, 2009).

[5] S. Tarucha et al., Phys. Rev. Lett. **77**, 3613 (1996).

[6] W. G. van der Wiel, S. De Franceschi, T. Fujisawa, J. M. Elzerman, S. Tarucha, L. P. Kouwenhoven, Science **289**, 2105 (2000).