Coulomb blockade and single electron tunneling

A typical semiconductor device utilizes many electron; for example there can be 10^{11} – 10^{12} electrons in 1 cm^2 area of a typical MOSFET device, so that even an area as small as 1 μm x 1 μm involve 10^3 – 10^4 electrons. On the other hand, if a device size well below 0.1 μm is achieved, then a single electron may be involved in the device application, where the concept of coulomb blockade plays an important role. This is based on the following observations: when an electron is transferred from lead to the small system (e.g., quantum dot) then there is a rearrangement of charge in the lead, resulting in a change in the electrostatic potential energy (see Fig. 8.1).

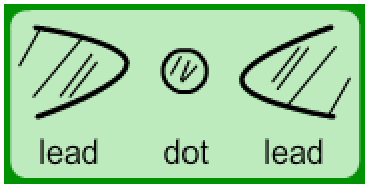

Fig. 8.1 : Schematic for single electron tunneling arrangement

For large system this charge is usually washed out by the thermal noise $\left(energy \sim k_B T\right)$, but for small systems, the change can be substantially larger than the thermal energy $k_B T$, especially at low temperature. Such large changes in the electrostatic energy due to transfer of a single charge results in a gap in the energy spectrum, which yields the so called Coulomb blockade (see Fig. 8.2).

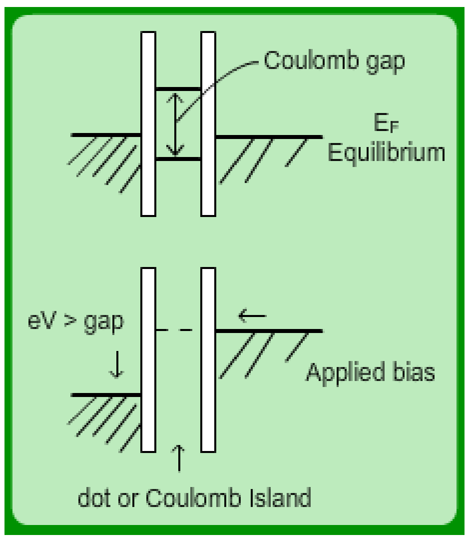

Fig. 8.2 : Schematic for energy gap in Coulomb blockade

To understand this phenomenon, it is convenient to consider an equivalent circuit of the dot (called Coulomb Island), as schematically shown in Fig. 8.3.

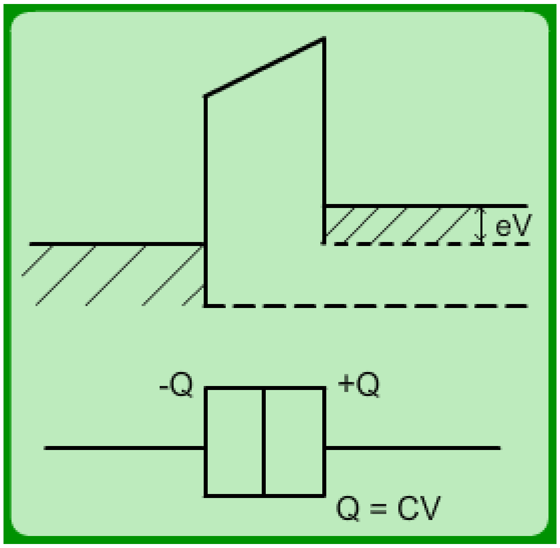

Fig. 8.3 : Schematic for Coulomb blockade

First consider a single tunnel ground biased by a volt V. An equivalent circuit for this can be drawn in terms of the junction capacitance.

The energy stored in the tunnel junction is $U = \frac{Q^2}{2C}$

In the presence of the bias V, an electron at the source electrode with K.E. $E_s(k)$ will tunnel into the state at the drain electrode with K.E. $E_d(k')$, so that energy conservation gives

$$E_s(k) + \frac{1}{2}CV^2 = E_d(k') + \frac{(CV - e)^2}{2C}$$

Since the electron tunneling through the junction must satisfy Pauli's exclusion principle, one has the following inequalities:

$$E_s(k) < E_F - k_B T \text{ and } E_d(k') > E_F + k_B T$$

These imply $E_d(k') - E_s(k) > 2k_B T$ so that the tunneling condition is given by

$$eV \geq \frac{e^2}{2C} - 2k_B T$$

and no current can flow at a bias below a threshold voltage which depends on the temperature. At T = 0, no current below V = e / (2C), which region is called the coulomb. The prohibition of tunneling is called coulomb blockade. The energy eV = e^2 / (2C) is called the coulomb gap energy and will be denoted by E_c.

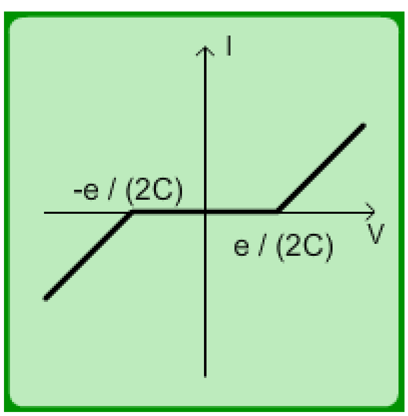

Fig. 8.4 : I-V characteristic illustrating Coulomb gap

This energy is usually very small for a junction of macroscopic size or is usually marked by the thermal noise making its detection impossible. However, in a tunneling junction with an area of 0.01 mm^2 with an insulating film of thickness 1nm, the energy E_c becomes equivalent to about 1K (Kelvin). Therefore if the junction is kept at temperature below 1K, then it is expected that the tunneling probability will be controlled by E_c. To illucidate the I-V characteristic, consider some data when the charges if ± Q are stored at the junction surfaces, the electrostatic energy of the system is $\frac{Q^2}{2C}$. When a single electron tunnels from say the negative electrode to the positive electrode the charges at the positive and negative electrodes will change from + Q to + (Q – e) and – Q to – (Q – e) respectively. As a result, the change in electrostatic energy of the system before and after one electron tunneling is given by

$$\Delta E = \frac{(Q-e)^2}{2C} - \frac{Q^2}{2C} = \frac{e}{C}\left(\frac{e}{2} - Q\right) = E_C - eV \quad \text{using V=Q/C.}$$

This tells us that as an electron tunnels through the junction, the system loses his own coulomb energy E_c and receives eV from the voltage source. Therefore, the tunneling condition for the electron in this system must satisfy the $V > E_C/e$. In other words, under the condition $V < E_C/e$ and at low temperature (such that $k_BT << E_C$ the electron is not allowed to tunnel through the junction. In a similar manner, one finds that the change in charges due to tunneling of an electron from positive electrode to negative electrode is given by $\pm(Q+e)$ so that the change in electrostatic energy is now given by $\Delta E = \frac{(Q+e)^2}{2C} - \frac{Q^2}{2C} = eV + E_C$

Therefore, for an applied voltage bias V such that $-E_C/e < V$, the electron is not allowed to tunnel through the junction.

Combining these results one finds that the tunneling in forbidden for $|V| < e$.

We now turn to the case of two junctions which are connected in series with an isolated area in between the coulomb island (the quantum dot). Such a system is sometimes called a Single Electron Transistor (SET). We can try to understand the characteristics of the SET in terms of the coulomb blockade. It is then simpler to consider its equivalent circuit (see Fig. 8.5).

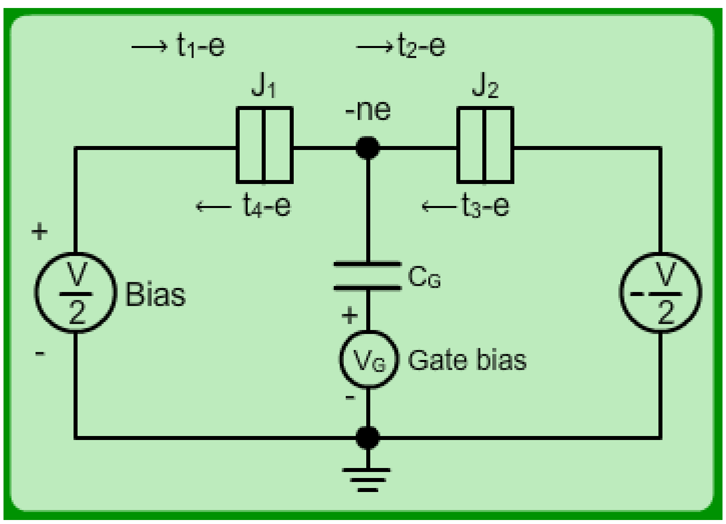

Fig. 8.5 : Equivalent circuit for a Single Electron Transistor

The electron number in the Coulomb Island is controlled externally through a junction capacitance C_G, and the capacitance is called the gate electrode or gate capacitance. In such a device the tunneling through the gate capacitance is usually neglected. To compare the device with a transistor, note that the contact leading to the left tunnel junction (J_1) is the source electrode and the contact outside the right tunnel junction (J_2) is the drain electrode. Consider now the situation where voltages = V/2 and – V/2 are applied to the source and drain electrodes respectively while the gate electrode is supplied independently with voltage V_G. Assume that there are negative charges in the Coulomb Island. Four tunneling processes t_1, t_2, t_3, t_4 are possible, where an electron is added or removed from n electrons (in the island) due to the tunneling. Let the change in the electrostatic energy due to the four tunneling processes be ΔE_1, ΔE_2, ΔE_3, ΔE_4 respectively. When all of them are positive, tunneling is forbidden and number of electron remains unchanged, giving rise to no current. For example the change in electrostatic energy ΔE_1 due to the tunneling process t_1 is

$$\Delta E_1 = \frac{e}{C_T}\left[C_T\frac{V}{2} - C_G V_G - e\left(n+\frac{1}{2}\right)\right], \qquad C_T = 2C + C_G$$

where C_T is the total electrostatic capacitance seen from the Coulomb island (it is practically assumed that the junction J_1 and J_2 are identified with same junction capacitance C).

Similarly $\Delta E_2, \Delta E_3, \Delta E_4$ are obtained by replacing in above the last term with $+\left(n+\frac{1}{2}\right), -\left(n-\frac{1}{2}\right)$ and $-\left(n+\frac{1}{2}\right)$ respectively.

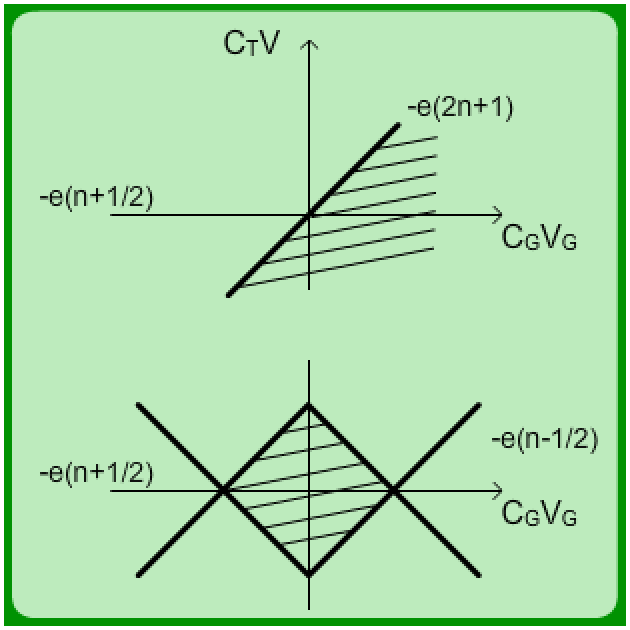

Fig. 8.6 : Illustration of Coulomb Diamond

The equation for ΔE_1 is usually depicted via C_TV against C_GV_G plot. The hatched region indicates the condition under which the tunneling t_1 cannot occur when n electron exist in the Coulomb island. The hatched region (square or rhombus) shows the condition when any of the processes t_1, t_2, t_3, t_4 cannot occur. This is usually called Coulomb Diamond (i.e., rhombus as in playing cards) (secc Fig. 8.6).

The condition for no tunneling is shown below for different n.

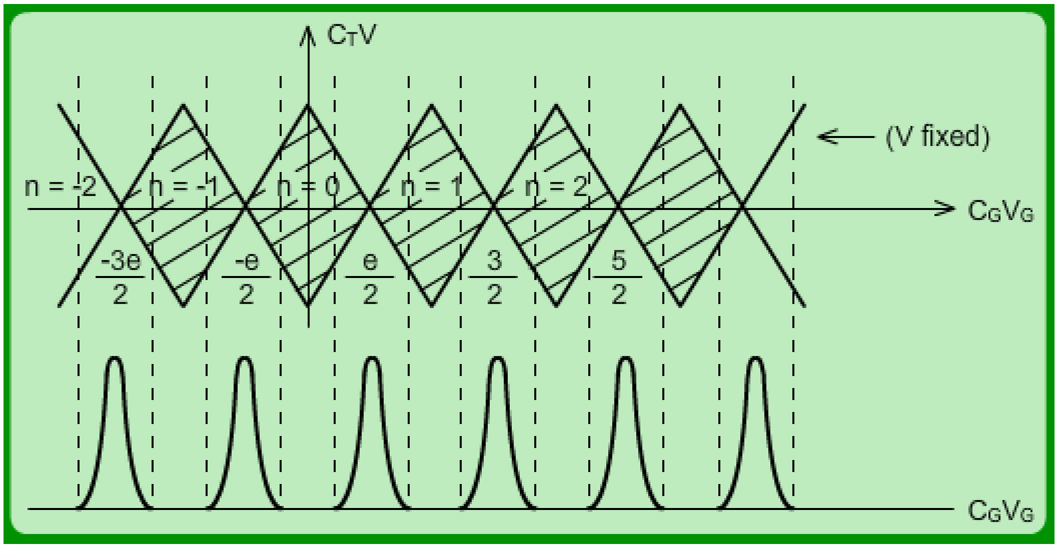

Fig. 8.7 : Schematic for Coulomb-Diamonds and the corresponding source-drain current as function of C_GV_G

If voltage applied to the source-drain electrodes are held fixed and the gate voltage V_G is varied, then the SET characteristics pass along a line as shown, which cuts through the diamonds of forbidden and allowed regions. As a result, the source-drain current flows periodically as shown (lower figure in Fig. 8.7). It is obvious that the change in the electronic charge at the gate electrode is less that e, the charge of one electron. It implies that the device is operated by a charge of less than a single election.

The term SET was named after the fact that such a device comprising of tunnel junction is operable by controlling the source-drain current by the gate voltage and the device features are comparable with those of the MOSFET or MESFET.

Note that to implement the Coulomb blockade effect to a real device at 300 K, one must have $e^2/2C$ much larger than k_BT for 300 K. At 300 K, this requires $C \leq 3.1 \times 10^{-18}$ F, a very very small capacitance, requiring the device to be very very small. SETs have been fabricated and their device operations have be confirmed by observing Coulomb blockade, even integrated SET's have been reported as SET memories. However, very large scale integration remains a challenge.

Reference

1. Transport in Nanostructures, D. K. Ferry and S. M. Goodwick, Cambridge Univ. Press. Cambridge, UK, 2001 Reprint, Ch. 4.
2. Physics of Semiconductor devices, J. P. Colinge and C. A. Colinge, Kluwer Academic Pub, 2002, Dordrech, Ch. 10.