Physics of Semiconductors (8)

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

June 15, 2016

2.4 *pn*-junction transistors

From left, John Bardeen, William Shockley, Walter Brattain. At AT&T Bell Laboratories, in 1948.

Today, we see two kinds of semiconductor devices invented by a genius named William Shockley. The style of research and development which he began, as well as his devices, has been changing the human life. The above expression is not exaggeration, I believe. I have read a short commentary, which tells "the researchers in Bell Labs. were doing basic research on the surface states of Ge with putting tips on the surfaces and accidentally found the transistor action". But this is far from real situation. Walter Brattain and John Bardeen, who were the direct finders, were doing research aiming at construction of "solid state amplifier" under the team leader Shockley. They did not expect such an easy finding probably but they realized the amplification certainly because they were doing such objective research.

The experiment was done a little before the Christmas of 1947 (said to be 12/16. The application for patent was 12/23) Shockley was out of the labs for a journey. He was thus not so glad hearing the success. Also the transistor (the term is a combination of transfer and resistor) which Brattain and Bardeen accidentally found was called "point contact type", unstable, had low reproducibility. It should have serious obstacles for commercial viability. Their finding might have stimulated Shockley's fight as an inventor, he was absorbed in thought as a theorist aiming at realization of "reproducible device for amplification" and finally got the brilliant inspiration of junction transistor, on the new year's eve allegedly. The theory for the junction transistor established 1/23 in the next year. The experimental realization was a year later. The event was the glorious dawn of the semiconductor physics, in which artificial structures in solids utilize the structural sensitivity of semiconductors and create new functions, new stages of physics[2].

2.4.1 Junction transistor: structure

Figure 2.4 shows basic structure of **junction transistor** (Bipolar Junction Transistor, BJT, at times just "bipolar transistor"), in which two pn-junctions are placed close to each other. npn and pnp are possible types of junctions. An ohmic contact to the central layer is required for the device to have three terminals. The terminals at the two ends are called **Collector** (C), **Emitter** (E) respectively and the central one is called **Base** (B). In the very beginning, the structure was fabricated with alloying metals which work as dopants to both sides of the base material. The naming "Base" came from the fact though lithography and thermal diffusion, ion implantation and epitaxy soon became the dominant methods. As we will see for the transistor action, the base should be very thin. Thinner than the minority carrier diffusion length.

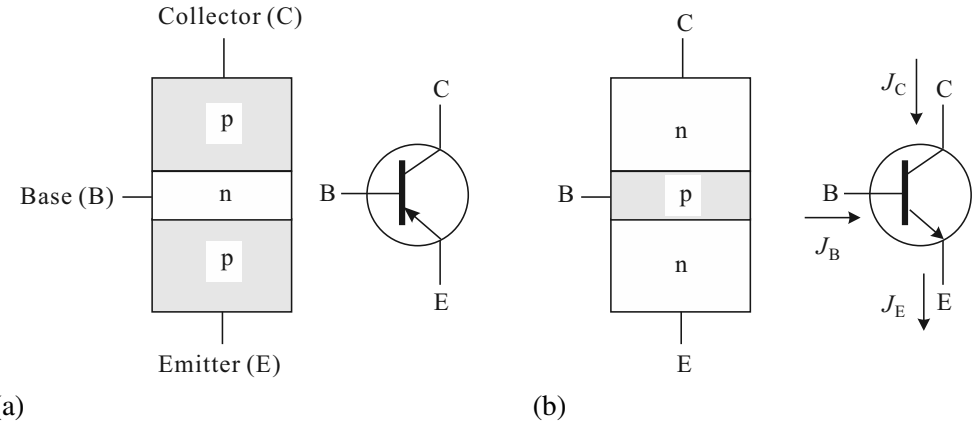


Figure 2.4: (a) Schematic structure of pnp transistor. Circuit symbol and the names of terminals. (b) Schematic structure of npn transistor and the circuit symbol.

Circuit symbols of transistors are shown in Fig.2.4, which represent connections of two electrodes to the base graphically. Circles are often omitted. pnp and npn are distinguished with the direction of arrow, which indicates direction of electric current when minority carriers are injected into the base electrode. Below we consider npn-type and define the directions of the current as in the figure.

2.4.2 Current-amplification of junction transistors

In the first experiment a constant voltage source is connected to B-C and collector current $J_{\rm C}$ is measured. Inside the structure B-C is nothing but a pn diode and the result is a well known rectification characteristics ($J_{\rm E} = 0$ in Fig.2.5(a)). Now we connect a constant current source between E and B, and apply finite currents through E. Because B-E is also a pn junction, the forward bias is positive for B. As shown in Fig.2.5(a) $V_{\rm BC} - J_{\rm C}$ curve shifts parallely to negative. The amount of shift is almost $J_{\rm E}$.

It should be noted that the characteristics is close to that of a solar cell shown in Fig.2.3(b). The similarity is not a coincidence, rather, the physical situation is almost the same. While In a solar cell, the minority carriers are directly created by photon irradiation, in a transistor, the minority carriers are injected through the pn junction between E and B to the other junction between B and C.

The phenomenon occurring in the junctions are summarised as follows. Here we only describe the phenomenon in conduction band while that in valence band can be discussed in parallel. In an npn junction, a reverse bias voltage to B(p)-C(n) suppresses the diffusion current from the n-layer to the p-layer. The (reverse) diffusion of *electrons* from the p-layer to the n-layer is not enhanced by the reverse bias because all the electrons reach from the p-layer to the junction are swung to the n-layer and it is already saturated at zero-bias. Under the reverse (or zero) bias condition of $V_{\rm BC}$, let the other pn-junction (E-B) be under a forward bias condition. This is possible because an Ohmic contact is attached to the base electrode, hence $V_{\rm EB}$ and $V_{\rm BC}$ can be controlled independently. The forward bias lowers the barrier by the built-in potential in E-B junction and the electrons (majority in the n-layer) diffuse into the base layer and the minority carrier concentration increases in B. This is the phenomenon called **minority carrier injection**, which decays over the **minority carrier diffusion length** through the recombination with majority carriers (holes). Note that the continuity in current is hold. The flow by injected electrons is not driven by the electric field but by the density gradient. So the flux is perpendicular to the junction plane almost ignoring the base Ohmic electrode (the recombination current goes to the electrode). When the B-layer is much thinner than the minority carrier diffusion length, most of the injected carriers reach the other junction enhancing the *reverse current*. In Fig.2.6(a), this appears as the enhancement of the reverse current, the amount of which is determined that of injected minority carriers. Hence the current does not depend on $V_{\rm BC}$ as long as there is no forward current.

Now an amplification circuit can be composed as follows. Let the electrodes C-E be voltage biased as in Fig.2.6(c). The amount of minority carrier injection into B-layer is determined by $V_{\rm BE}$. Hence in this circuit

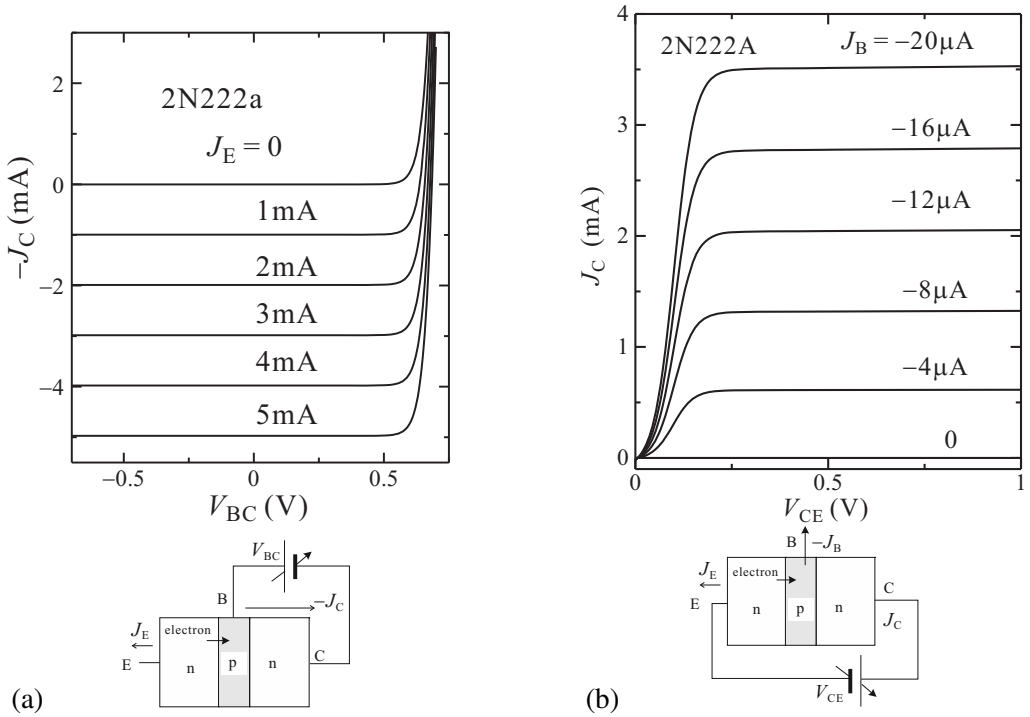


Figure 2.5: (a) $J_{\rm C}$ (upside down for convenience) as a function of $V_{\rm BC}$ with $J_{\rm E}$ as a parameter in the circuit shown in the lower panel. With increasing $J_{\rm E}$ *i.e.*, injecting electrons from E to B, the characteristics resembles to that of an illuminated solar cell. (b) Application of collector-emitter voltage $V_{\rm CE}$ with floating B, almost no current flows due to the reverse bias in C-B. The biasing B with some currents $J_{\rm C}$ appears according to $J_{\rm B}$ showing saturation for $V_{\rm CE}$.

$J_{\rm C}$ strongly depends on $V_{\rm BE}$ as shown in Fig.2.6(a). However, the relation is too non-linear for the use of the device in a voltage-input circuit.

Some of the injected minority carriers recombine with majority carriers and some portion flows out to B-electrode. The base current $J_{\rm B}$ depends on $V_{\rm BE}$ in the same functional form only but the coefficient as $J_{\rm C}$ because the pn-junction is the same. $J_{\rm C}$ is thus proportional $J_{\rm B}$, that is,

$$J_{\rm C} = h_{\rm FE} J_{\rm B}. \tag{2.16}$$

The good linearity is confirmed in the measurement as shown in Fig.2.6(b). $h_{\rm FE}$ is called **current amplification factor**. And it is often said that "a bipolar transistor works as a current amplification device" from this face. This is in practice, true as long as we use it as a black box device in electric circuits. However in physical mechanism, as discussed above, there is no such causality that a small current drives a larger current. The following expression may be closer to reality: a small current here is just a monitor for voltage to control a large current.

In the usage of a BJT in a circuit, care should be taken that (because it is a "current amplification device") the input voltage bias should be set to a low differential resistance region. Particularly in high frequency circuits, the impedance matching should be taken to the characteristic impedance of the transmission line. One simple "rule" for transistor circuits is that when a transistor is working as an amplifier, the base-emitter bias voltage should be around the quasi-threshold voltage (though as we saw there is no threshold voltage in pn-junctions, in ordinary circuit scale, the I-V curve seems as if it has).

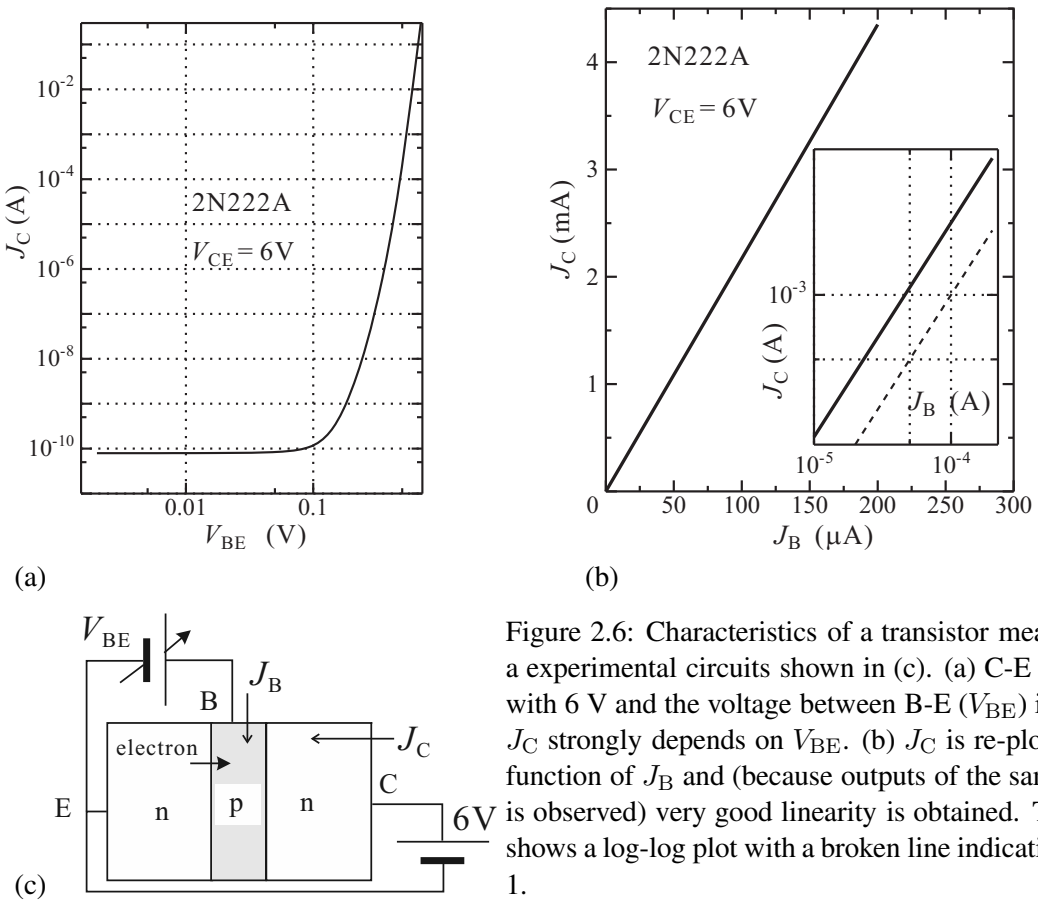


Figure 2.6: Characteristics of a transistor measured in a experimental circuits shown in (c). (a) C-E is biased with 6 V and the voltage between B-E ($V_{\rm BE}$) is varied. $J_{\rm C}$ strongly depends on $V_{\rm BE}$. (b) $J_{\rm C}$ is re-plotted as a function of $J_{\rm B}$ and (because outputs of the same diode is observed) very good linearity is obtained. The inset shows a log-log plot with a broken line indicating slope 1.

2.5 Field effect transistors I

Field Effect Transistors (FETs) are now used much more widely in circuits than BJTs. And the idea of FET was born even long before that of BJT [1], but for the realization of FET requires technologies even higher than those for BJT and the realization was later than that for BJT. In these 20 years, Metal-Oxide-Semiconductor (**MOS**) type FETs are mainly used but the first FET was realized for Junction FET (JFET), which utilize pn junctions.

2.5.1 pn-junction and depletion layer

For understanding the device action of JFET, the relation between the reverse bias voltage and the depletion layer is important. We consider a pn-junction shown in Fig.2.7, with x-dependent potential $\phi(x)$. The Poisson equation is given as

$$\frac{d^2\phi}{dx^2} = -aq(x) \quad (a \equiv (\epsilon\epsilon_0)^{-1}). \tag{2.17}$$

In the space-charge region (depletion layer) we assume abrupt concentration distribution of dopants and sharp cutting of the end of depletion layer. Then

$$\begin{cases} q = -eN_A & (-w_p \le x \le 0), \\ q = eN_D & (0 \le x \le w_n). \end{cases} \tag{2.18}$$

[1] Shockley wrote a patent on FET before BJT though many similar ideas had existed before that. We cannot say the patent is as unique as that of BJT.

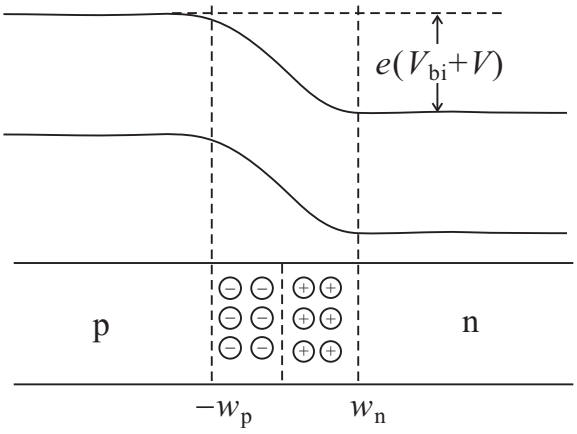


Figure 2.7: Simple model of a pn junction

Let's take the asymptotic condition as $\phi(-\infty) = 0$. When there is external reverse bias voltage V, the boundary condition at the edges of depletion layer is

$$\begin{aligned} \phi(-w_p) = 0, \quad & \left.\frac{d\phi}{dx}\right|_{-w_p} = 0, \\ \phi(w_n) = V + V_{\rm bi}, \quad & \left.\frac{d\phi}{dx}\right|_{w_n} = 0. \end{aligned} \tag{2.19}$$

Integration of the above gives

$$\phi(x) = \begin{cases} (aeN_A/2)(x+w_p)^2 & (-w_p \le x \le 0), \\ V + V_{\rm bi} - (aeN_D/2)(x-w_n)^2 & (0 \le x \le w_n). \end{cases} \tag{2.20}$$

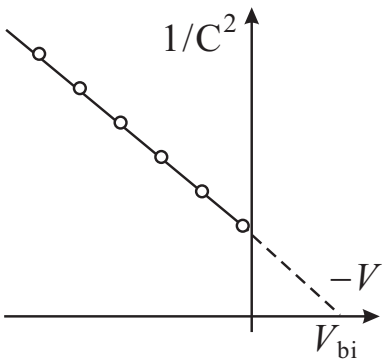
From the condition for the connection at $x = 0$

$$\lim_{x\to+0}\phi = \lim_{x\to-0}\phi, \qquad \lim_{x\to+0}(d\phi/dx) = \lim_{x\to-0}(d\phi/dx), \tag{2.21}$$

the widths of depletion layer w_p, w_n are given as follows.

$$w_p = \left[\frac{2\epsilon_0\epsilon(V+V_{\rm bi})}{eN_A}\cdot\frac{N_D}{N_D+N_A}\right]^{1/2}, \quad w_n = \left[\frac{2\epsilon_0\epsilon(V+V_{\rm bi})}{eN_D}\cdot\frac{N_A}{N_D+N_A}\right]^{1/2} \tag{2.22}$$

$$w_d = w_p + w_n = \left[\frac{2\epsilon_0\epsilon(V+V_{\rm bi})}{e}\cdot\frac{N_A+N_D}{N_AN_D}\right]^{1/2}. \tag{2.23}$$



The charge accumulated in the depletion layer on n-side is $Q = eN_Dw_d$ per unit area giving the effective capacitance (differential capacitance) as

$$\frac{dQ}{dV} = eN_D\sqrt{\frac{2\epsilon\epsilon_0}{eN_D}}\frac{1}{2\sqrt{V+V_{\rm bi}}} = \sqrt{\frac{\epsilon\epsilon_0eN_D}{2}}(V+V_{\rm bi})^{-1/2}. \tag{2.24}$$

In a p^+n-structure, that is, $N_A \gg N_D$,

$$w_d \approx \left[\frac{2\epsilon\epsilon_0(V+V_{\rm bi})}{eN_D}\right]^{1/2} \approx w_n. \tag{2.25}$$

This means the depletion layer expands in proportional to the square root of the reverse bias voltage plus the built-in potential.

This relationship is frequently used for characterization of pn-junctions. For example, differential capacitance $C(V)$ can be measured with applying high frequency voltage

source with a small amplitude and through the phase shift. We plot the data as shown in the left figure (for the convenience, the horizontal axis is taken to $-V$), $1/C^2$ versus $-V$. If $N_{\rm D}$ is spatially uniform, the data points should be aligned on a line. (2.24) is valid only for $V > 0$ and $C \to \infty$ cannot be realized. But with extrapolation from $V > 0$ the point $1/C^2 = 0$ can be specified and we obtain $V_{\rm bi}$ from this.

When N_D is not uniform spatially or some deep level traps exist, we obtain information of the spatial distribution from differentiating the plot. Application of pulses in V and analysis of transient response under light illumination or related techniques can bring much of the information inside the semiconductor[4].

2.5.2 Junction Field Effect Transistors

Figure 2.8 shows a schematic drawing of the JFET structure in a cross sectional view. It is for an n-channel, which has two electrodes on the both edges. They are called **Source** (S) and **Drain** (D) respectively. The channel is sandwiched by p^+ layers called **Gates** (G).

The principle of device action is very simple as can be seen in Fig.2.8. Applying reverse bias to the gates causes expansion of white-colored depletion layer according to eq.(2.23). This makes the conduction channel narrower and enhances the channel resistance up to infinity for pinch-off. Thus the current through the device is controlled by the gate voltage. This is apparently a voltage-controlled device and the input impedance is typically resistance of pn-junction in reverse bias condition. So it is classified into high input impedance device.

A characteristic feature here is that a large source-drain current causes a significant voltage drop across the device, resulting in gradient of effective reverse bias voltage for the channel-controlling depletion layer. Let us see a simple model. As before in the model for pn-junctions, we assume the boundaries between depletion layers and conduction channel are abrupt. Let the gate length L, the thickness of JFET $2w_t$. We take the channel direction along y-axis. The depletion layer with w_d is

$$w_d(y) = \sqrt{\frac{2\epsilon\epsilon_0 V(y)}{eN_D}}, \tag{2.26}$$

where $V(y)$ is local voltage at position y between the channel and the gate. $V(y)$ can be obtained by subtracting voltage along the channel $V_{\rm ch}$ due to the source-drain current from the sum of the built-in potential $V_{\rm bi}$ and the reverse bias gate voltage $V_{\rm g}$.

$$V(y) = V_{\rm g} + V_{\rm bi} - V_{\rm ch}(y).$$

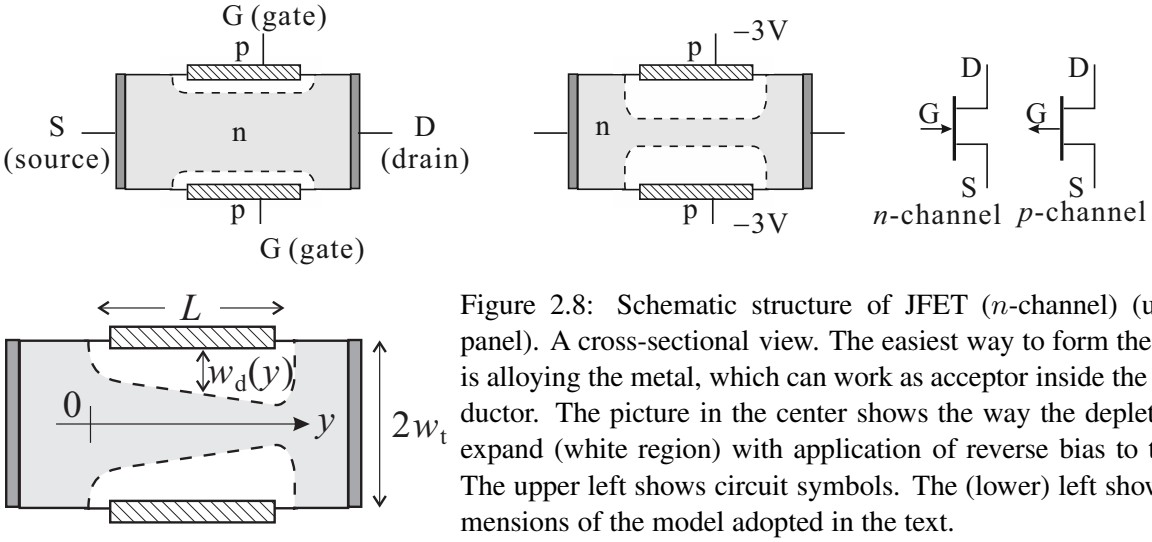


Figure 2.8: Schematic structure of JFET (n-channel) (upper left panel). A cross-sectional view. The easiest way to form the p^+-layer is alloying the metal, which can work as acceptor inside the semiconductor. The picture in the center shows the way the depletion layer expand (white region) with application of reverse bias to the gates. The upper left shows circuit symbols. The (lower) left shows the dimensions of the model adopted in the text.

We have no injection of minority carrier and only consider the drift current of majority carriers. The electric field along y-direction is dV/dy. Let the channel depth W and the drift current through the channel is

$$J_{ch} = eN_D\mu_n \frac{dV}{dy} \cdot 2(w_t - w_d)W. \tag{2.27}$$

In steady state there is no charging up and $J_{\rm ch}$ is uniform through the channel thus integration over the channel should be $J_{\rm ch}L$.

$$J_{\rm ch}L = \int_0^L J_{\rm ch}dy = 2eN_D\mu_n W \int_0^L (w_t - w_d)\frac{dV}{dy}dy = 2w_t eN_D\mu_n W \int_{V_0}^{V_L} \left(1 - \frac{w_d}{w_t}\right) dV. \tag{2.28}$$

Let the critical voltage $V_{\rm c}$ at which the channel is pinched ($w_d = w_t$) and $J_{ch} = 0$ then $V_c = eN_D w_t^2/2\epsilon\epsilon_0$. Hence from $w_d/w_t = \sqrt{V/V_c}$, $J_{\rm ch}$ in this model is obtained as

$$J_{\rm ch} = \frac{2N_D e\mu_n W w_t}{L}\left[V_L - V_0 + \frac{2}{3\sqrt{V_c}}(V(V_0)^{3/2} - V(V_L)^{3/2})\right]. \tag{2.29}$$

In eq.(2.29), at small voltages, the first linear term in V_L is dominant and $J_{\rm ch}$ increases linearly. With increasing the voltage, the last $V_L^{3/2}$ term grows and at last the current begins decreasing, which means negative differential resistance. In actual device, this does not occur and $J_{\rm ch}$ simply saturates with increasing V. The model contains various shortages, *e.g.*, the equipotential lines are straight and along x-axis. Improved models can reproduce the saturation but they are inevitably complicated. There are also empirical analytical formulas well fit to the experiments but they have no physical reasoning.

2.6 Field Effect Transistors II

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

2.6.1 Schottky barrier (junction)

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

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

On semiconductor surfaces, there usually are **surface states** with high density of states. Metal-semiconductor junctions are strongly affected by those states. Here, however, we first look what Anderson's rule tells about the interface[5]. The baseline of rigid bands can be taken to an edge of "band", in which electrons can freely travel between the metal and the semiconductor. It is usually impossible to find such an energy band inside insulators and semiconductors, which have very different energy bands. Then such a "band" can be found as

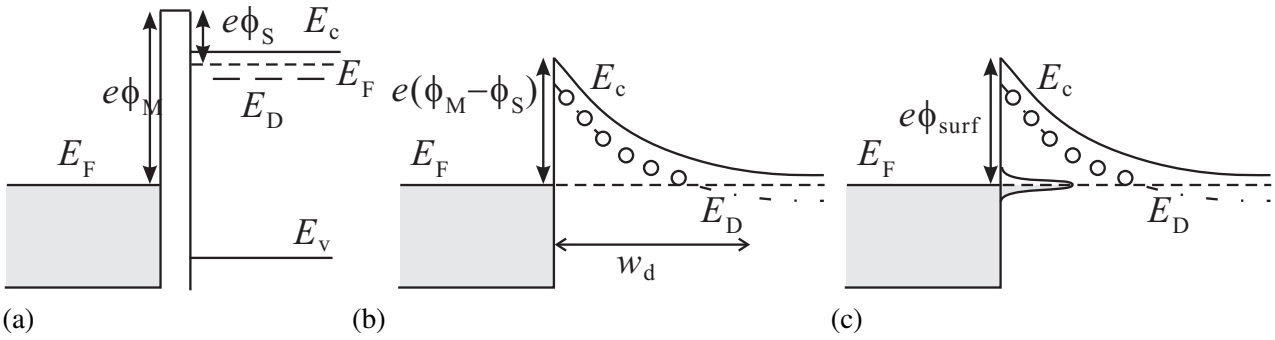


Figure 2.9: (a) Virtual band alignment, in which a metal and a semiconductor are connected as the vacuum levels for them agree. (b) Band bending effect to make $E_{\rm F}$ constant throughout the junction is superposed to the alignment in (a). The situation corresponds to an ideal interface without surface states at the semiconductor side. (c) Illustration of Fermi level pinning by surface states. The surface potential $\phi_{\rm surf}$ is determined by the position of the dominant surface states from the band edge $E_{\rm c}$. This usually has nothing to do with the difference between the work function.

the vacuum levels. Then the excitation energy required is so called **work function**. Let the work functions in the semiconductor and the metal $e\phi_{\rm S}$ and $e\phi_{\rm M}$ respectively. Generally $e\phi_{\rm M} \neq e\phi_{\rm S}$. On the other hand, from the guiding principle 2., the bulk $E_{\rm F}$'s in the metal and in the semiconductor away from the junction should be the same. And $E_{\rm F}$ should be constant throughout the system.

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

We assume $e\phi_{\rm M}$ is larger than $e\phi_{\rm S}$, the semiconductor is doped to n-type and the donor concentration is $N_{\rm D}$. We make the vacuum levels in the both sides fit to each other and extrapolate the bulk band structures to the interface to obtain the band alignment shown in Fig.2.9(a). Here the Fermi level in the semiconductor places higher than that in the metal causing flow of carriers from the semiconductor to the metal. The carrier flow generates charge accumulation at the interface creating an electric field perpendicular to the junction plane. The metallic side is also charged up but it has much higher charge concentration, which screens the electric field within the screening length less than a lattice constant making the band bending negligible in this side. Let the accumulated charge in the metal side per unit area $-Q$, in the semiconductor side ($x > 0$, interface at $x = 0$), the electric field at x is $(eN_D x - Q)/\epsilon\epsilon_0$ and the potential difference between 0 and x_d is

$$\phi(x_d) = \int_0^{x_d} (eN_D x - Q)/\epsilon\epsilon_0 dx = \frac{1}{\epsilon\epsilon_0}\left(\frac{eN_D}{2}x_d^2 - Qx_d\right). \tag{2.30}$$

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

$$Q = \sqrt{2\epsilon\epsilon_0 N_D e(\phi_M - \phi_S)}, \quad \therefore w_d = \sqrt{\frac{2\epsilon\epsilon_0(\phi_M - \phi_S)}{eN_D}} \equiv \sqrt{\frac{2\epsilon\epsilon_0 V_s}{eN_D}}. \tag{2.31}$$

Here we write $eV_s \equiv \phi_M - \phi_S$. Now we can illustrate the band structure for electrons (holes for p-type) around the metal-semiconductor interface as in Fig.2.9(b), showing a potential barrier, which is called **Schottky barrier**.

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

equation for thermal electron emission from metallic surface can be applied to obtain

$$J = AT^2 \left[\exp\left(\frac{e(V - V_s)}{k_{\rm B}T} \right) - \exp\left(\frac{-eV_s}{k_{\rm B}T} \right) \right] = eAT^2 \exp\left(\frac{-eV_s}{k_{\rm B}T} \right) \left[\exp\left(\frac{eV}{k_{\rm B}T} \right) - 1 \right]. \qquad (2.32)$$

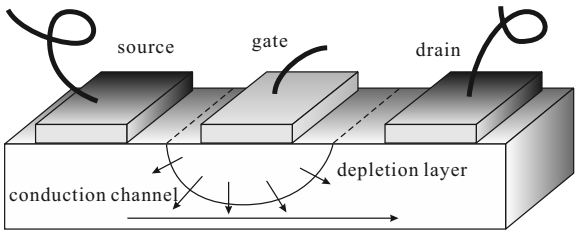
Here A is the **Richardson coefficient**. The first term is current from the semiconductor side, the second is that from the metal side. The current-voltage characteristics is similar to that of a pn-junction with the Schttkey barrier height corresponding to the built-in potential.

In the above the surface of semiconductor is too much idealized for it to have no surface states. However in real metal-semiconductor junctions, current-voltage characteristics are similar to eq.(2.32). One big difference is in eq.(2.32), the barrier height should change with changing the metal species but in reality, the barrier height is almost constant for semiconductor species and independent of metals. This is due to the **surface states** on the semiconductors. The surface states have narrow energy widths, very high density of states pinning the Fermi level to the center of them. Hence the band bending exists even before the connection to metals and the alignment is accomplished between the metal $E_{\rm F}$ and the surface states. This is called **pinning of Fermi level** by the surface states.

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

2.6.2 MES-FET

Among III-V semiconductors, GaAs is frequently used for electric devices as well as for optical devices. But it is difficult to form good quality oxide layers on the surfaces, hence no MOS type device for GaAs is available. Instead, MEtal-Semiconductor FET (MES-FET) structure has been frequently adopted. GaAs has light electron mass, high mobilities. And the effective capacitance of Schottky diode can be small. Hence GaAs MESFETs are often used for high-frequency application.



As shown in the left figure, the structure of MES-FET is simple. The conduction channel thickness is controlled with the reverse bias voltage (**gate voltage**) through that of depletion layer. The device action, characteristics are similar to those for JFET. Schottky junctions have larger leak current in gate characteristics, only single carrier type is available and complementary circuits cannot be composed with them. These properties are great obstacles for large scale integration. MES-FETs are still widely used as high frequency devices for *e.g.*, microwave.

2.6.3 MOS structure

As named, a thin oxide film for insulation is inserted between a metal and a semiconductor in a Metal-Oxide-Semiconductor (MOS) structure. Needless to say, most frequently used Si has SiO_2 as the oxide layer, which is very stable and has good insulation characteristics. An SiO_2 film can be easily formed with thermal oxidation onto a Si. Both p-type and n-type channels can be controlled and Complementary MOS (CMOS) circuits are easily realized. Also with low gate leakage current, high on-conductance, off-resistance, the power consumption in logic circuits jumped down with the CMOS circuits hence increased degree of integration. Now CMOS is doubtlessly the king of semiconductor circuits. A few decades ago high speed logic circuits were

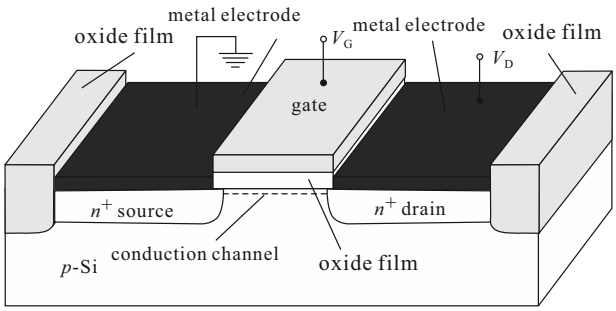


Figure 2.10: Schematic view of a MOSFET device. In fabrication holes are opened on thermally oxidized films with lithography. The dopants are diffused through the holes. The structure like this often appears due to the process.

mainly composed with Emitter Coulpled Logic (ECL) of BJT but the requirement of large scale integration and the increase of cut-off frequency in CMOS circuit have made drastic change and now, even so called supercomputers are using CMOS circuit in CPU.

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

References

[1] S. M. Sze, K. K. Ng, "Physics of semiconductor devices", (Wiley-Blackwell, 2007).

[2] Jon Gertner, "The Idea Factory: Bell Labs and the Great Age of American Innovation", (Penguin Press, 2012).

[3] Lectures on experimental physics "Basic measurement techniques" (Maruzen, 1999) Ch.2 (in Japanese)

[4] A. Patanè and N. Balkan eds. *Semiconductor Research: Experimental Techniques* (Springer Series in Materials Science, 2012).

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

Appendix C：Deep level transient spectroscopy (DLTS)

Here I would like to give qualitative explanation on the basic principles of Deep Level Transient Spectroscopy (DLTS). For details, see *e.g.* ref. [4]. We consider modification to effective capacitance (2.24), which depends on the reverse bias voltage V. Let $N_{\rm D}$ be the shallow donor concentration, $N_{\rm P}$ the one for a deep donor. In the region where this deep donor responds to change in the bias voltage, the voltage-differential capacitance is expressed as a function of reverse voltage V as

$$w_d(V) \approx \left[\frac{2\epsilon\epsilon_0(V+V_{\rm bi})}{e(N_{\rm D}+N_{\rm P})}\right]^{1/2} \approx w_n, \tag{C.1}$$

$$C(V) = \sqrt{\frac{\epsilon\epsilon_0 e(N_{\rm D}+N_{\rm P})}{2}}(V+V_{\rm bi})^{-1/2}. \tag{C.2}$$

For simplicity, we consider the situation that the reverse bias V_p is applied and kept for sufficiently long time for electrons to escape from the depletion layer including the deep levels [2]. Now V is abruptly lowered to

[2] At low temperatures the capture/emission rates of deep levels become very small and it is not rare that we need days for the emission. So this condition is, in general, hard to be fulfilled. But the consideration of this does not give significant change and thus we adopt the assumption.

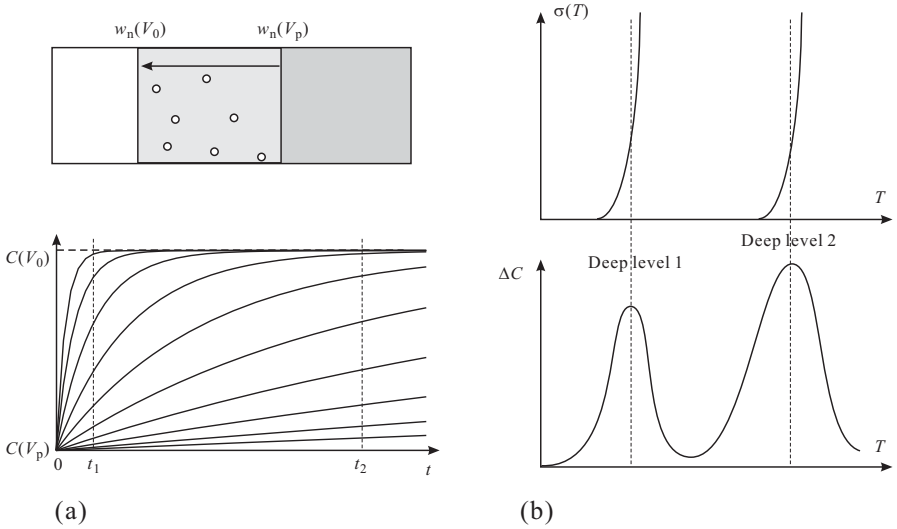


Figure 2.11: (a) Upper panel: Illustration that the change in the reverse bias $V_p \to V_0$ makes shallow levels and a part of deep levels ready for catching carriers. Lower panel: With progress in capture of carriers, differential capacitance $C(V)$ shows transient response. (b) Upper panel: two deep levels exist and assumed temperature dependences of the capture cross section σ are illustrated. Lower panel: shows how the DLTS signal appears from the temperature dependence $\sigma(T)$.

$V_0 < V_p$ and the carriers are captured by the donor levels within $w(V_0) < x \leq w(V_p)$. Shallow donors have high capture rate and can respond within ms without delay, deep levels, on the other hand, the capture rate strongly depends on temperature and with decreasing temperature, the average time for capture often elongates from ms to s, min, hour and sometimes day. Then if we open up a fixed time window and observe the time evolution of C, the time dependence is observed in the time window at some temperature range and in low or high temperature regions the effect of deep levels does not observed.

Such a process is illustrated in Fig.2.11(a). We take $t = 0$ at the time the reverse bias is changed:$V_p \to V_0$ and measure the difference in the differential capacitances at t_1 and t_2: $\Delta C = |C(t_1) - C(t_2)|$ as a function of temperature T.

We now assume existence of two species of deep donors, which have temperature dependent capture cross sections shown in the upper panel of Fig.2.11(b). ΔC should show two peaks in the temperature dependence. Analysis of the data gives the concentration and capture cross section of each deep level, and combination with photo-response, in some cases identification of deep levels or at least energy positions can be measured[4]. With variation of V_0 and V_p, depth profile of deep levels can be obtained also.