The Junction Diode

Basic Operation

The diode is fabricated of a semiconductor material, usually silicon, which is doped with two impurities. One side is doped with a donor or n-type impurity which releases electrons into the semiconductor lattice. These electrons are not bound and are free to move about. Because there is no net charge in the donor impurity, the n-type semiconductor is electrically neutral. The other side is doped with an acceptor or p-type impurity which imparts free holes into the lattice. A hole is the absence of an electron which acts as a positive charge. The p-type semiconductor is also electrically neutral because the acceptor material adds no net charge.

Figure 1(a) illustrates the cross section of the diode. The junction is the dividing line between the n-type and p-type sides. Thermal energy causes the electrons and holes to move randomly. Electrons diffuse across the junction into the p-type side and holes diffuse across the junction into the n-type side. This causes a net positive charge to develop in the n-type side and a net negative charge to develop in the p-type side. These charges set up an electric field across the junction which is directed from the n-type side to the p-type side. The electric field opposes further diffusion of the electrons and holes. The region in which the electric field exists is called the depletion region. There are no free electrons or holes in this region because the electric field sweeps them out.

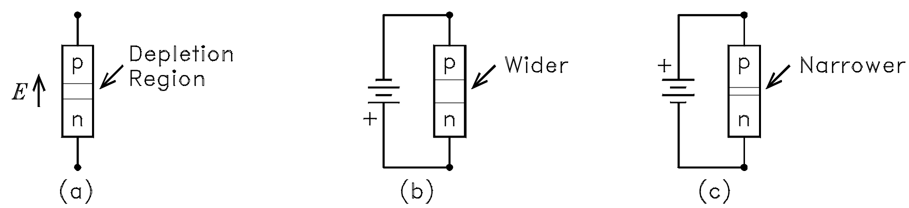

Figure 1: (a) Diode cross section. (b) Reverse biased diode. (c) Forward biased diode.

Figure 1(b) shows the diode with a battery connected across it. The polarity of the battery is such that it reinforces the electric field across the junction causing the depletion region to widen. The positive terminal pulls electrons in the n-type side away from the junction. The negative terminal pulls holes in the p-type side away from the junction. No current can flow. The diode is said to be reverse biased. Figure 1(c) shows the diode with the battery polarity reversed. The battery now tends to cancel out the electric field in the depletion region, causing its width to decrease. The positive terminal forces holes toward the junction. The negative terminal forces electrons toward the junction. A current flows which increases rapidly if the applied voltage is increased. The diode is said to be forward biased.

$i - v$ Characteristics

Figure 2(a) shows the circuit symbol for the diode. The arrow part of the symbol points in the direction of current flow when the diode is forward biased. The upper terminal is called the anode. The lower terminal is called the cathode. These names come from vacuum tube diodes.

The theoretical equation for the diode current is

$$i_D = I_S \left[\exp\left(\frac{v_D}{nV_T}\right) - 1 \right]$$

where I_S is the saturation current, n is the emission coefficient, and V_T is the thermal voltage. The emission coefficient accounts for recombinations of electrons and holes in the depletion region, which tend to decrease the current. For discrete diodes, it has the value $n \simeq 2$. For integrated circuit diodes, it has the value $n \simeq 1$. The reason it is different for the two cases is because an integrated circuit diode is fabricated as a bipolar transistor with the collector connected to the base. The impurity doping in transistors is done so as to minimize recombinations. Thus, $n \simeq 1$ when recombinations can be neglected.

A typical plot of i_D versus v_D is given in Fig. 2(b). For $v_D \leq 0.6\,\text{V}$, the current is very small. For $v_D > 0.6\,\text{V}$, the current increases rapidly with v_D. The voltage at which the diode appears to begin conducting is called the cutin voltage. This is approximately 0.6 V for the plot in Fig. 2(b).

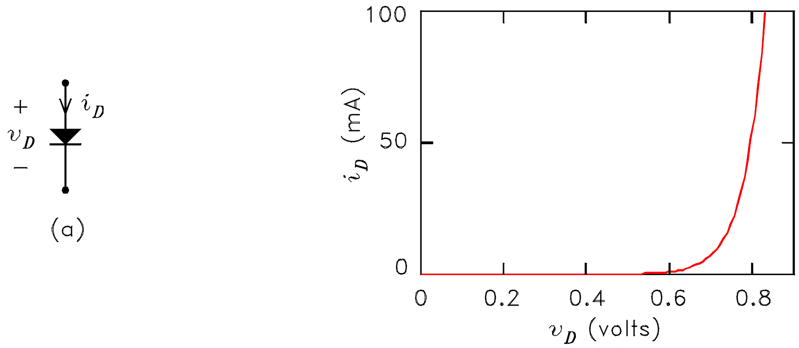

Figure 2: (a) Diode symbol. (b) Typical current versus voltage.

The thermal voltage is given by

$$V_T = \frac{kT}{q}$$

where k is Boltzmann's constant, T is the Kelvin temperature, and q is the electronic charge. At $T = 290\,\mathrm{K}$, the thermal voltage has the value $V_T = 0.025\,\mathrm{V}$. The default value for T in SPICE is $T = 300\,\mathrm{K}$. In this case, the thermal voltage has the value $V_T = 0.02585\,\mathrm{V}$. This value should be used in any hand calculations that are to be compared to SPICE simulations. Because V_T increases with T, the equation for i_D seems to imply that increasing T decreases i_D. However, I_S increases rapidly with temperature which causes i_D to increase with T. A rule of thumb that is often quoted for silicon diodes is that if i_D is held constant, v_D decreases by about $2\,\mathrm{mV}$ for each degree C as T increases.

Example 1 *Calculate the amount by which* v_D *would have to be increased to double the diode current. Assume* $\exp(v_D/nV_T) \gg 1$ *and* $V_T = 0.025\,\mathrm{V}$*.*

Solution. We have

$$2 = \frac{I_S\left[\exp\left(v_{D2}/nV_T\right) - 1\right]}{I_S\left[\exp\left(v_{D1}/nV_T\right) - 1\right]} \simeq \exp\left(\frac{v_{D2} - v_{D1}}{nV_T}\right)$$

Solution for $\Delta v = v_{D2} - v_{D1}$ yields

$$\begin{aligned} \Delta v &= nV_T \ln 2 \\ &= 17.3\,\mathrm{mV} \text{ for } n = 1 \\ &= 34.7\,\mathrm{mV} \text{ for } n = 2 \end{aligned}$$

This example illustrates how fast the diode current increases with increasing diode voltage.

Reverse Breakdown

If the diode is reverse biased and the voltage is increased, a point will be reached when the diode enters reverse breakdown and a current will flow. The voltage at which this occurs is called the zener voltage. Fig. 3 illustrates the variation of diode current with voltage. The reverse breakdown voltage is labeled $-V_Z$ and the cutin voltage is labeled V_γ. In diode applications, the reverse breakdown voltage must be greater than the maximum applied reverse voltage to prevent diode failure. As an example, diodes used to rectify $120\,\mathrm{V}$ rms ac line voltage must have a reverse breakdown voltage greater than $120\sqrt{2} = 170\,\mathrm{V}$. Diodes which are fabricated to have a specific value of V_Z are called zener diodes. These diodes are used as voltage reference diodes. The zener or reverse breakdown voltage is a function of the n and p type doping levels in the diode. The higher the doping, the lower the value of V_Z. For V_Z less than approximately $4\,\mathrm{V}$, the breakdown is due to zener breakdown. For V_Z greater than approximately $5\,\mathrm{V}$, the breakdown is due to avalanche breakdown.

Figure 3: Diode characteristics showing the zener voltage V_Z and the cutin voltage V_γ.

An example application of a zener diode is shown in Fig. 4. The circuit shows a dc source having a voltage V_1 connected to a zener diode and a load through a resistor R_1. If V_1 changes, V_L remains approximately constant. Thus the diode acts as a constant voltage reference. The diode current is

$$I_Z = \frac{V_1 - V_Z}{R_1} - I_L$$

The power dissipated in the diode is given by $V_Z I_Z$. The power dissipated in the resistor R_1 is $(V_1 - V_Z)(I_Z + I_L)$. The power dissipated in the diode decreases as I_L increases and the power dissipated in R_1 increases as I_L increases. The power ratings of the diode and R_1 must be high enough to prevent thermal failure under worst case conditions for each. This means that the power rating of the diode must be chosen for the minimum anticipated value of I_L and the power rating of R_1 must be chosen for the maximum anticipated value of I_L.

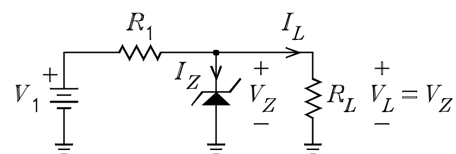

Figure 4: Example zener diode regulator.

Example 2 *For* $V_1 = 24\,\text{V}$*,* $V_Z = 12\,\text{V}$*, and* $2\,\text{mA} \le I_L \le 10\,\text{mA}$*, calculate the value of* R_1 *such that* $I_Z \ge 2\,\text{mA}$ *and determine the minimum power ratings of* R_1 *and the zener diode in Fig. 4.*

Solution. The value of R_1 is calculated for the maximum value of I_L. Thus the maximum current through R_1 is $2\,\text{mA}+10\,\text{mA} = 12\,\text{mA}$. Thus $R_1 = (24\,\text{V} - 12\,\text{V})/12\,\text{mA} = 1\,\text{k}\Omega$. The minimum power rating for R_1 is $12\,\text{V}\times 12\,\text{mA} = 144\,\text{mW}$. The minimum power rating for the zener diode is $12\,\text{V}\times(12\,\text{mA} - 2\,\text{mA}) = 100\,\text{mW}$.

Linear Model

The linear model of the diode approximates the $i - v$ characteristics by a straight line that is tangent to the actual curve at the dc bias point. Fig. 5(a) shows the curve with the tangent line at the point (V_D, I_D). The curve intersects the horizontal axis at the voltage V_{DO}. For small changes in v_D and i_D about the tangent point, the tangent line gives a good approximation to the actual curve. The slope of the tangent line is given by

$$m = \left.\frac{di_D}{dv_D}\right|_{(V_D, I_D)} = \frac{1}{nV_T} I_S \exp\left(\frac{V_D}{nV_T}\right) = \frac{I_D + I_S}{nV_T} = \frac{1}{r_d}$$

where the units of r_d are ohms. The equation of the tangent line is

$$i_D = \frac{V_D - V_{DO}}{r_d}$$

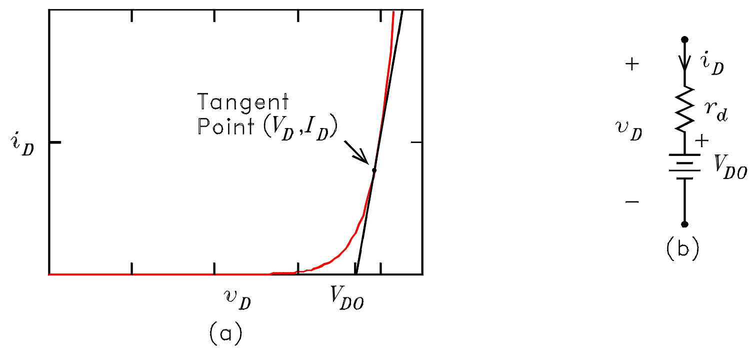

Figure 5: (a) $i - v$ characteristics with tangent line at (V_D, I_D). (b) Linear diode model.

The equivalent circuit which models this equation is shown in Fig. 5(b). Because $v_D = V_D$ when $i_D = I_D$, it follows that the voltage V_{DO} is given by

$$V_{DO} = V_D - I_D r_d$$

Example 3 *A diode is biased at* $I_D = 1\,\mathrm{mA}$. *It is given that* $I_S = 3.6 \times 10^{-9}\,\mathrm{A}$, $n = 2$, *and* $V_T = 25.9\,\mathrm{mV}$. *Solve for* r_d *and* V_{DO} *in the linear model.*

Solution. $V_D = nV_T \ln(I_D/I_S) = 0.649\,\mathrm{V}$, $r_d = nV_T/I_D = 51.8\,\Omega$, $V_{DO} = V_D - I_D r_d = 0.598\,\mathrm{V}$.

Small-Signal Model

Because the diode equation for i_D as a function of v_D is non-linear, the tools of linear circuit analysis cannot be applied, in general, to circuits containing diodes. However, if the diode current is known for a particular voltage, linear circuit analysis can be used to predict the change in current for a given change in voltage, provided the change is not very large. Such an approach is called a small-signal analysis.

Let the diode voltage and current be written

$$v_D = V_D + v_d \qquad i_D = I_D + i_d$$

where V_D and I_D are dc bias values and v_d and i_d are small-signal changes about the bias values. Let the diode equation be denoted by $i_D = f(v_D)$. We can write

$$I_D + i_d = f(V_D + v_d) \simeq f(V_D) + f'(V_D)\, v_d$$

where the approximation is a first-order Taylor series expansion about the point (V_D, I_D). Because $I_D = f(V_D)$, we can solve for i_d to obtain

$$i_d = f'(V_D)\, v_d = \frac{d}{dV_D}\left\{I_S\left[\exp\left(\frac{V_D}{nV_T}\right) - 1\right]\right\} \times v_d = \frac{I_S}{V_T}\exp\left(\frac{V_D}{nV_T}\right) \times v_d = \frac{I_D + I_S}{nV_T} \times v_d$$

The small-signal resistance is defined as the ratio of v_d to i_d and is given by

$$r_d = \frac{nV_T}{I_D + I_S} \simeq \frac{nV_T}{I_D}$$

This is the same r_d as in the linear model of the diode in Fig. 5(b). This equation says that the diode small-signal resistance is inversely proportional to the current through it. Each time the current is doubled, the resistance is halved. It follows from the linear diode model that r_d can be interpreted graphically as the reciprocal of the slope of the i_D versus v_D curve at the point (V_D, I_D).

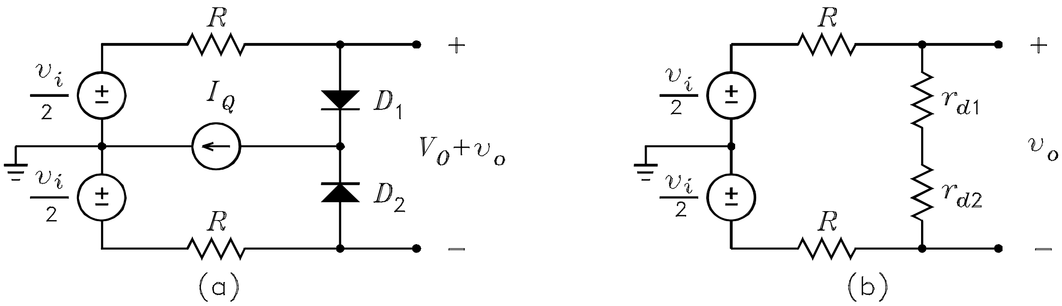

Figure 6: (a) Diode attenuator circuit. (b) Small-signal circuit for calculating v_o.

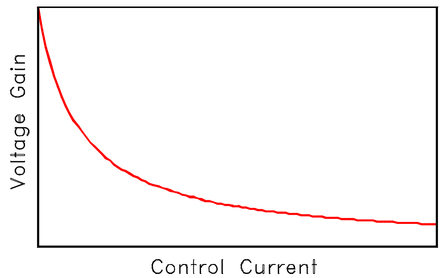

Figure 7: Plot of voltage gain versus control current.

Example 4 *Figure 6(a) shows a diode attenuator circuit. The input voltage v_i is a small-signal differential voltage represented by two series sources with the common terminal grounded. The current I_Q is a dc control current. Solve for the small-signal gain v_o/v_i. Assume the diodes are identical, $n = 2$, and $V_T = 0.025\,\mathrm{V}$.*

Solution. For the dc solution, we set $v_i = 0$. In this case, $v_o = 0$. For identical diodes, the current I_Q splits equally between the two. It follows the dc voltages across the diodes cancel in calculating V_O so that $V_O = 0$. The small-signal resistance of each diode is $r_{d1} = r_{d2} = 2nV_T/I_Q = 0.1/I_Q$. The small-signal circuit is shown in Fig. 6(b). The current I_Q does not appear in this circuit because it is a dc source which becomes an open circuit when zeroed. It follows by voltage division that

$$\frac{v_o}{v_i} = \frac{r_{d1} + r_{d2}}{2R + r_{d1} + r_{d2}} = \frac{2 \times 0.1/I_Q}{2R + 2 \times 0.1/I_Q} = \frac{1}{1 + 10 I_Q R}$$

Thus the small-signal gain of the circuit can be varied by varying the dc current I_Q. A typical plot of the gain versus current is shown in Fig. 7.