that ionization is the mechanism whereby an electron can absorb sufficient energy to break away from the atomic structure and enter the conduction band. You will note that the energy associated with each electron is measured in *electron volts* (eV). The unit of measure is appropriate, since

$$\boxed{W = QV} \quad \text{eV} \tag{1.2}$$

as derived from the defining equation for voltage $V = W/Q$. The charge Q is the charge associated with a single electron.

Substituting the charge of an electron and a potential difference of 1 volt into Eq. (1.2) will result in an energy level referred to as one *electron volt.* Since energy is also measured in joules and the charge of one electron $= 1.6 \times 10^{-19}$ coulomb,

$$W = QV = (1.6 \times 10^{-19}\ \text{C})(1\ \text{V})$$

and

$$\boxed{1\ \text{eV} = 1.6 \times 10^{-19}\ \text{J}} \tag{1.3}$$

At 0 K or absolute zero (−273.15°C), all the valence electrons of semiconductor materials find themselves locked in their outermost shell of the atom with energy levels associated with the valence band of Fig. 1.8b. However, at room temperature (300 K, 25°C) a large number of valence electrons have acquired sufficient energy to leave the valence band, cross the energy gap defined by E_g in Fig. 1.8b and enter the conduction band. For silicon E_g is 1.1 eV, for germanium 0.67 eV, and for gallium arsenide 1.41 eV. The obviously lower E_g for germanium accounts for the increased number of carriers in that material as compared to silicon at room temperature. Note for the insulator that the energy gap is typically 5 eV or more, which severely limits the number of electrons that can enter the conduction band at room temperature. The conductor has electrons in the conduction band even at 0 K. Quite obviously, therefore, at room temperature there are more than enough free carriers to sustain a heavy flow of charge, or current.

We will find in Section 1.5 that if certain impurities are added to the intrinsic semiconductor materials, energy states in the forbidden bands will occur which will cause a net reduction in E_g for both semiconductor materials—consequently, increased carrier density in the conduction band at room temperature!

1.5 EXTRINSIC MATERIALS—*n*- AND *p*-TYPE

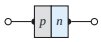
The characteristics of semiconductor materials can be altered significantly by the addition of certain impurity atoms into the relatively pure semiconductor material. These impurities, although only added to perhaps 1 part in 10 million, can alter the band structure sufficiently to totally change the electrical properties of the material.

A semiconductor material that has been subjected to the doping process is called an extrinsic material.

There are two extrinsic materials of immeasurable importance to semiconductor device fabrication: *n*-type and *p*-type. Each will be described in some detail in the following paragraphs.

n-Type Material

Both the *n*- and *p*-type materials are formed by adding a predetermined number of impurity atoms into a germanium or silicon base. The *n*-type is created by introducing those impurity elements that have *five* valence electrons (*pentavalent*), such as *antimony, arsenic,* and *phosphorus.* The effect of such impurity elements is indicated in



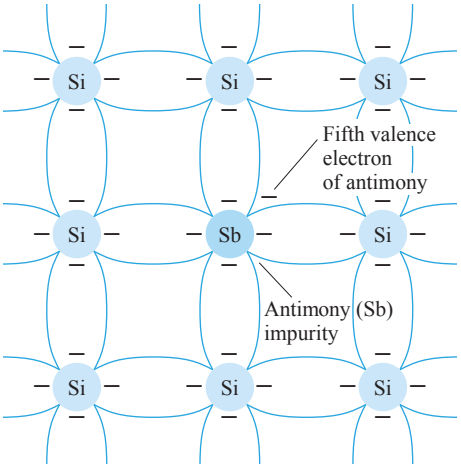


Figure 1.9 Antimony impurity in *n*-type material.

Fig. 1.9 (using antimony as the impurity in a silicon base). Note that the four covalent bonds are still present. There is, however, an additional fifth electron due to the impurity atom, which is *unassociated* with any particular covalent bond. This remaining electron, loosely bound to its parent (antimony) atom, is relatively free to move within the newly formed *n*-type material. Since the inserted impurity atom has donated a relatively "free" electron to the structure:

Diffused impurities with five valence electrons are called donor atoms.

It is important to realize that even though a large number of "free" carriers have been established in the *n*-type material, it is still electrically *neutral* since ideally the number of positively charged protons in the nuclei is still equal to the number of "free" and orbiting negatively charged electrons in the structure.

The effect of this doping process on the relative conductivity can best be described through the use of the energy-band diagram of Fig. 1.10. Note that a discrete energy level (called the *donor level*) appears in the forbidden band with an E_g significantly less than that of the intrinsic material. Those "free" electrons due to the added impurity sit at this energy level and have less difficulty absorbing a sufficient measure of thermal energy to move into the conduction band at room temperature. The result is that at room temperature, there are a large number of carriers (electrons) in the conduction level and the conductivity of the material increases significantly. At room temperature in an intrinsic Si material there is about one free electron for every 10^{12} atoms (1 to 10^9 for Ge). If our dosage level were 1 in 10 million (10^7), the ratio ($10^{12}/10^7 = 10^5$) would indicate that the carrier concentration has increased by a ratio of 100,000:1.

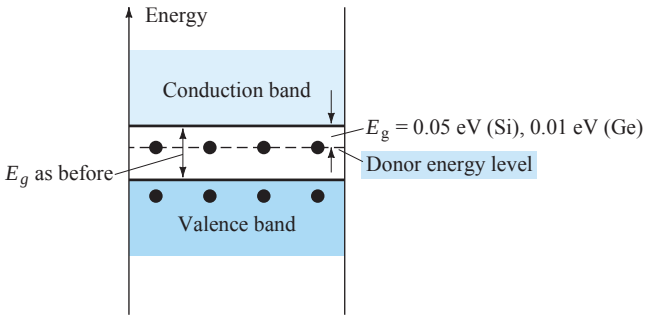


Figure 1.10 Effect of donor impurities on the energy band structure.

p-Type Material

The *p*-type material is formed by doping a pure germanium or silicon crystal with impurity atoms having *three* valence electrons. The elements most frequently used for this purpose are *boron, gallium,* and *indium.* The effect of one of these elements, boron, on a base of silicon is indicated in Fig. 1.11.

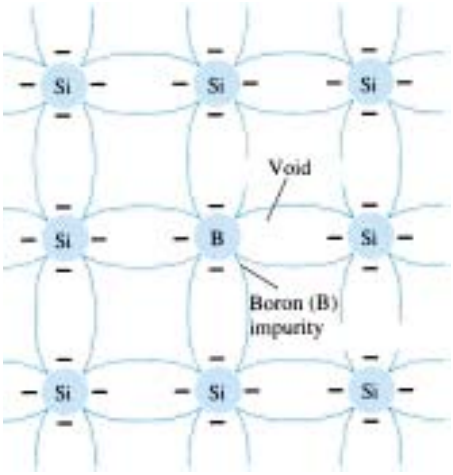


Figure 1.11 Boron impurity in *p*-type material.

Note that there is now an insufficient number of electrons to complete the covalent bonds of the newly formed lattice. The resulting vacancy is called a *hole* and is represented by a small circle or positive sign due to the absence of a negative charge. Since the resulting vacancy will readily *accept* a "free" electron:

The diffused impurities with three valence electrons are called acceptor atoms.

The resulting *p*-type material is electrically neutral, for the same reasons described for the *n*-type material.

Electron versus Hole Flow

The effect of the hole on conduction is shown in Fig. 1.12. If a valence electron acquires sufficient kinetic energy to break its covalent bond and fills the void created by a hole, then a vacancy, or hole, will be created in the covalent bond that released the electron. There is, therefore, a transfer of holes to the left and electrons to the right, as shown in Fig. 1.12. The direction to be used in this text is that of *conventional flow,* which is indicated by the direction of hole flow.

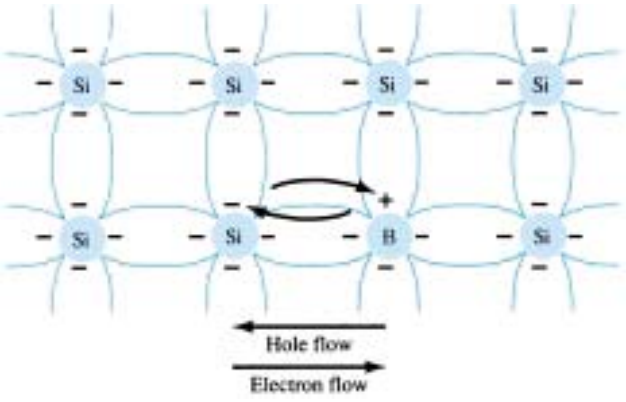


Figure 1.12 Electron versus hole flow.

Majority and Minority Carriers

In the intrinsic state, the number of free electrons in Ge or Si is due only to those few electrons in the valence band that have acquired sufficient energy from thermal or light sources to break the covalent bond or to the few impurities that could not be removed. The vacancies left behind in the covalent bonding structure represent our very limited supply of holes. In an *n*-type material, the number of holes has not changed significantly from this intrinsic level. The net result, therefore, is that the number of electrons far outweighs the number of holes. For this reason:

In an n-type material (Fig. 1.13a) the electron is called the majority carrier and the hole the minority carrier.

For the *p*-type material the number of holes far outweighs the number of electrons, as shown in Fig. 1.13b. Therefore:

In a p-type material the hole is the majority carrier and the electron is the minority carrier.

When the fifth electron of a donor atom leaves the parent atom, the atom remaining acquires a net positive charge: hence the positive sign in the donor-ion representation. For similar reasons, the negative sign appears in the acceptor ion.

The *n*- and *p*-type materials represent the basic building blocks of semiconductor devices. We will find in the next section that the "joining" of a single *n*-type material with a *p*-type material will result in a semiconductor element of considerable importance in electronic systems.

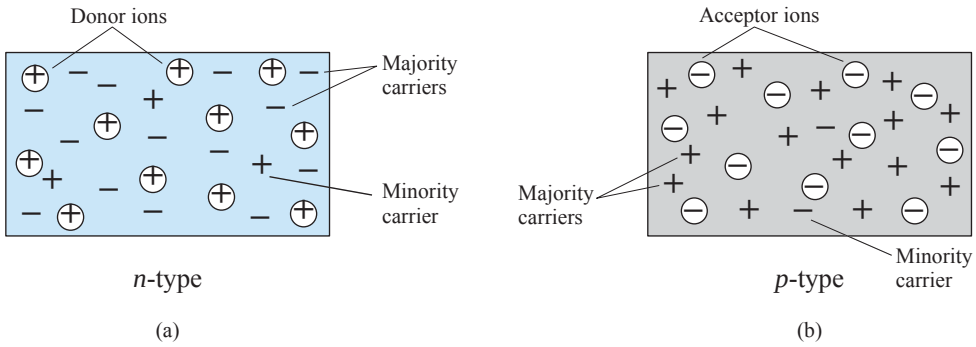


Figure 1.13 (a) *n*-type material; (b) *p*-type material.

1.6 SEMICONDUCTOR DIODE

In Section 1.5 both the *n*- and *p*-type materials were introduced. The semiconductor diode is formed by simply bringing these materials together (constructed from the same base—Ge or Si), as shown in Fig. 1.14, using techniques to be described in Chapter 20. At the instant the two materials are "joined" the electrons and holes in the region of the junction will combine, resulting in a lack of carriers in the region near the junction.

This region of uncovered positive and negative ions is called the depletion region due to the depletion of carriers in this region.

Since the diode is a two-terminal device, the application of a voltage across its terminals leaves three possibilities: *no bias* ($V_D = 0$ V), *forward bias* ($V_D > 0$ V), and *reverse bias* ($V_D < 0$ V). Each is a condition that will result in a response that the user must clearly understand if the device is to be applied effectively.

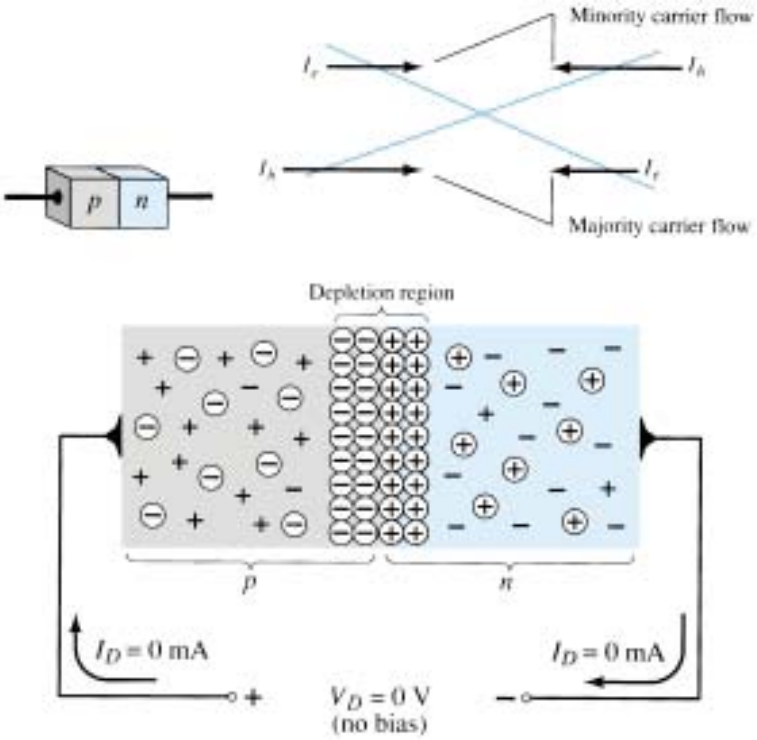


Figure 1.14 *p-n* junction with no external bias.

No Applied Bias ($V_D = 0$ V)

Under no-bias (no applied voltage) conditions, any minority carriers (holes) in the *n*-type material that find themselves within the depletion region will pass directly into the *p*-type material. The closer the minority carrier is to the junction, the greater the attraction for the layer of negative ions and the less the opposition of the positive ions in the depletion region of the *n*-type material. For the purposes of future discussions we shall assume that all the minority carriers of the *n*-type material that find themselves in the depletion region due to their random motion will pass directly into the *p*-type material. Similar discussion can be applied to the minority carriers (electrons) of the *p*-type material. This carrier flow has been indicated in Fig. 1.14 for the minority carriers of each material.

The majority carriers (electrons) of the *n*-type material must overcome the attractive forces of the layer of positive ions in the *n*-type material and the shield of negative ions in the *p*-type material to migrate into the area beyond the depletion region of the *p*-type material. However, the number of majority carriers is so large in the *n*-type material that there will invariably be a small number of majority carriers with sufficient kinetic energy to pass through the depletion region into the *p*-type material. Again, the same type of discussion can be applied to the majority carriers (holes) of the *p*-type material. The resulting flow due to the majority carriers is also shown in Fig. 1.14.

A close examination of Fig. 1.14 will reveal that the relative magnitudes of the flow vectors are such that the net flow in either direction is zero. This cancellation of vectors has been indicated by crossed lines. The length of the vector representing hole flow has been drawn longer than that for electron flow to demonstrate that the magnitude of each need not be the same for cancellation and that the doping levels for each material may result in an unequal carrier flow of holes and electrons. In summary, therefore:

In the absence of an applied bias voltage, the net flow of charge in any one direction for a semiconductor diode is zero.

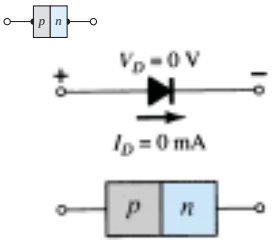


Figure 1.15 No-bias conditions for a semiconductor diode.

The symbol for a diode is repeated in Fig. 1.15 with the associated *n*- and *p*-type regions. Note that the arrow is associated with the *p*-type component and the bar with the *n*-type region. As indicated, for $V_D = 0$ V, the current in any direction is 0 mA.

Reverse-Bias Condition ($V_D < 0$ V)

If an external potential of *V* volts is applied across the *p*-*n* junction such that the positive terminal is connected to the *n*-type material and the negative terminal is connected to the *p*-type material as shown in Fig. 1.16, the number of uncovered positive ions in the depletion region of the *n*-type material will increase due to the large number of "free" electrons drawn to the positive potential of the applied voltage. For similar reasons, the number of uncovered negative ions will increase in the *p*-type material. The net effect, therefore, is a widening of the depletion region. This widening of the depletion region will establish too great a barrier for the majority carriers to overcome, effectively reducing the majority carrier flow to zero as shown in Fig. 1.16.

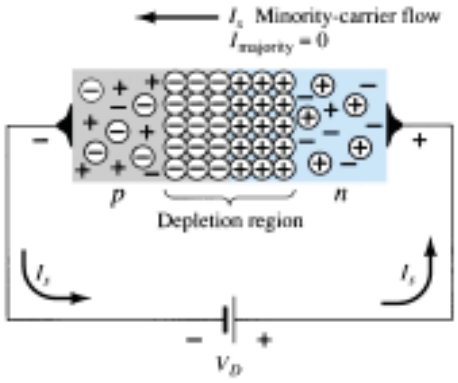


Figure 1.16 Reverse-biased *p*-*n* junction.

The number of minority carriers, however, that find themselves entering the depletion region will not change, resulting in minority-carrier flow vectors of the same magnitude indicated in Fig. 1.14 with no applied voltage.

The current that exists under reverse-bias conditions is called the reverse saturation current and is represented by I_s.

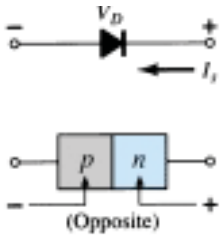


Figure 1.17 Reverse-bias conditions for a semiconductor diode.

The reverse saturation current is seldom more than a few microamperes except for high-power devices. In fact, in recent years its level is typically in the nanoampere range for silicon devices and in the low-microampere range for germanium. The term *saturation* comes from the fact that it reaches its maximum level quickly and does not change significantly with increase in the reverse-bias potential, as shown on the diode characteristics of Fig. 1.19 for $V_D < 0$ V. The reverse-biased conditions are depicted in Fig. 1.17 for the diode symbol and *p*-*n* junction. Note, in particular, that the direction of I_s is against the arrow of the symbol. Note also that the *n*egative potential is connected to the *p*-type material and the *p*ositive potential to the *n*-type material—the difference in underlined letters for each region revealing a reverse-bias condition.

Forward-Bias Condition ($V_D > 0$ V)

A *forward-bias* or "on" condition is established by applying the positive potential to the *p*-type material and the negative potential to the *n*-type material as shown in Fig. 1.18. For future reference, therefore:

A semiconductor diode is forward-biased when the association p*-type and positive and* n*-type and negative has been established.*

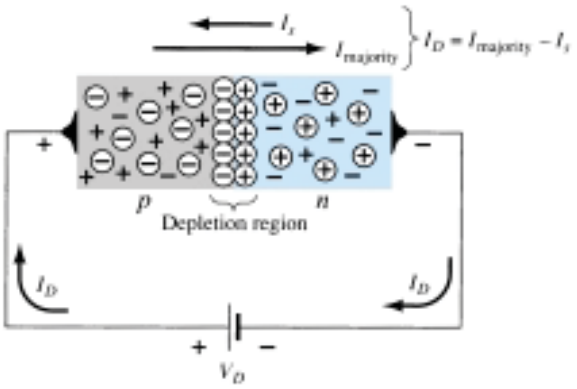


Figure 1.18 Forward-biased *p-n* junction.

The application of a forward-bias potential V_D will "pressure" electrons in the *n*-type material and holes in the *p*-type material to recombine with the ions near the boundary and reduce the width of the depletion region as shown in Fig. 1.18. The resulting minority-carrier flow of electrons from the *p*-type material to the *n*-type material (and of holes from the *n*-type material to the *p*-type material) has not changed in magnitude (since the conduction level is controlled primarily by the limited number of impurities in the material), but the reduction in the width of the depletion region has resulted in a heavy majority flow across the junction. An electron of the *n*-type material now "sees" a reduced barrier at the junction due to the reduced depletion region and a strong attraction for the positive potential applied to the *p*-type material. As the applied bias increases in magnitude the depletion region will continue to decrease in width until a flood of electrons can pass through the junction, re-

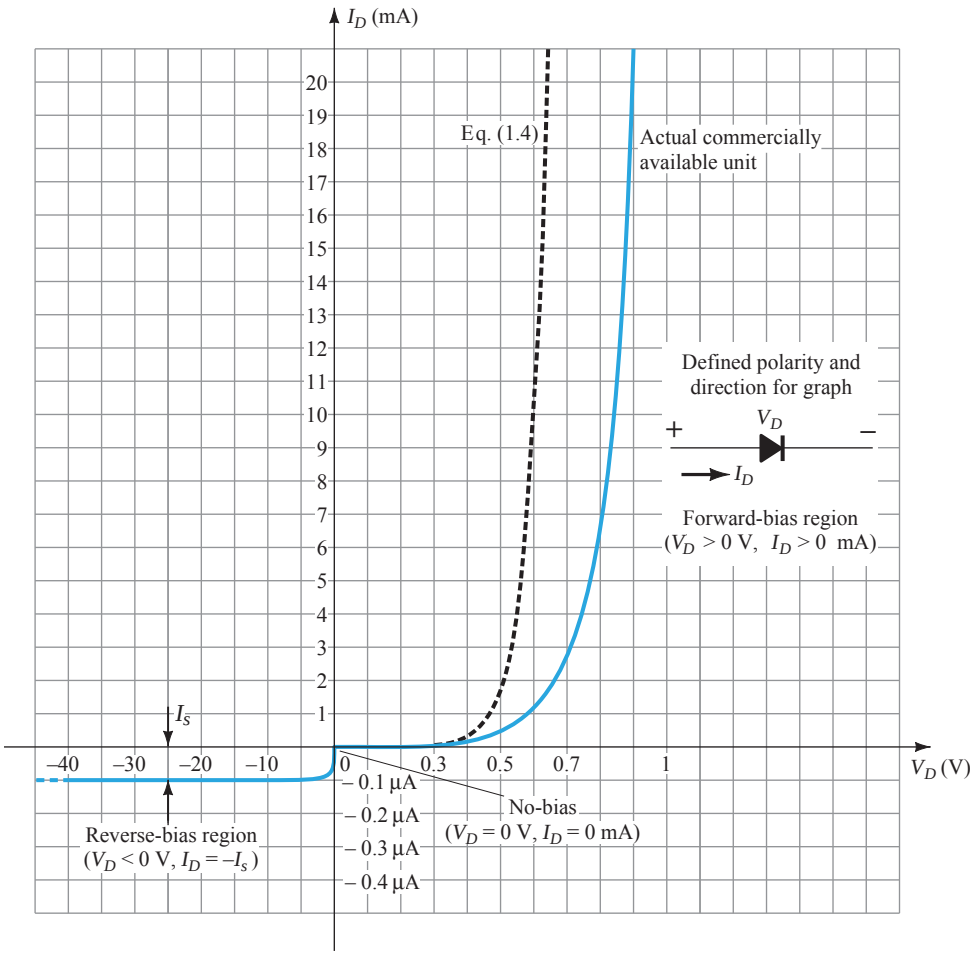
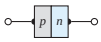


Figure 1.19 Silicon semiconductor diode characteristics.



sulting in an exponential rise in current as shown in the forward-bias region of the characteristics of Fig. 1.19. Note that the vertical scale of Fig. 1.19 is measured in milliamperes (although some semiconductor diodes will have a vertical scale measured in amperes) and the horizontal scale in the forward-bias region has a maximum of 1 V. Typically, therefore, the voltage across a forward-biased diode will be less than 1 V. Note also, how quickly the current rises beyond the knee of the curve.

It can be demonstrated through the use of solid-state physics that the general characteristics of a semiconductor diode can be defined by the following equation for the forward- and reverse-bias regions:

$$I_D = I_s(e^{kV_D/T_K} - 1) \tag{1.4}$$

where I_s = reverse saturation current

$k = 11{,}600/\eta$ with $\eta = 1$ for Ge and $\eta = 2$ for Si for relatively low levels of diode current (at or below the knee of the curve) and $\eta = 1$ for Ge and Si for higher levels of diode current (in the rapidly increasing section of the curve)

$T_K = T_C + 273°$

A plot of Eq. (1.4) is provided in Fig. 1.19. If we expand Eq. (1.4) into the following form, the contributing component for each region of Fig. 1.19 can easily be described:

$$I_D = I_s e^{kV_D/T_K} - I_s$$

For positive values of V_D the first term of the equation above will grow very quickly and overpower the effect of the second term. The result is that for positive values of V_D, I_D will be positive and grow as the function $y = e^x$ appearing in Fig. 1.20. At $V_D = 0$ V, Eq. (1.4) becomes $I_D = I_s(e^0 - 1) = I_s(1 - 1) = 0$ mA as appearing in Fig. 1.19. For negative values of V_D the first term will quickly drop off below I_s, resulting in $I_D = -I_s$, which is simply the horizontal line of Fig. 1.19. The break in the characteristics at $V_D = 0$ V is simply due to the dramatic change in scale from mA to μA.

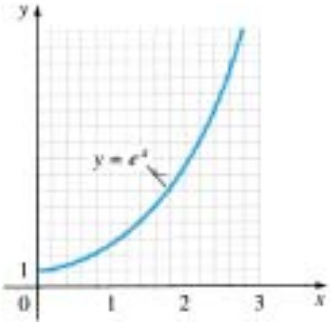


Figure 1.20 Plot of e^x.

Note in Fig. 1.19 that the commercially available unit has characteristics that are shifted to the right by a few tenths of a volt. This is due to the internal "body" resistance and external "contact" resistance of a diode. Each contributes to an additional voltage at the same current level as determined by Ohm's law ($V = IR$). In time, as production methods improve, this difference will decrease and the actual characteristics approach those of Eq. (1.4).

It is important to note the change in scale for the vertical and horizontal axes. For positive values of I_D the scale is in milliamperes and the current scale below the axis is in microamperes (or possibly nanoamperes). For V_D the scale for positive values is in tenths of volts and for negative values the scale is in tens of volts.

Initially, Eq. (1.4) does appear somewhat complex and may develop an unwarranted fear that it will be applied for all the diode applications to follow. Fortunately, however, a number of approximations will be made in a later section that will negate the need to apply Eq. (1.4) and provide a solution with a minimum of mathematical difficulty.

Before leaving the subject of the forward-bias state the conditions for conduction (the "on" state) are repeated in Fig. 1.21 with the required biasing polarities and the resulting direction of majority-carrier flow. Note in particular how the direction of conduction matches the arrow in the symbol (as revealed for the ideal diode).

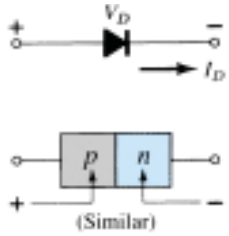


Figure 1.21 Forward-bias conditions for a semiconductor diode.

Zener Region

Even though the scale of Fig. 1.19 is in tens of volts in the negative region, there is a point where the application of too negative a voltage will result in a sharp change

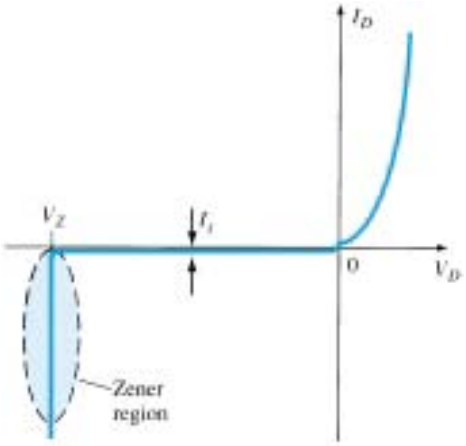


Figure 1.22 Zener region.

in the characteristics, as shown in Fig. 1.22. The current increases at a very rapid rate in a direction opposite to that of the positive voltage region. The reverse-bias potential that results in this dramatic change in characteristics is called the *Zener potential* and is given the symbol V_Z.

As the voltage across the diode increases in the reverse-bias region, the velocity of the minority carriers responsible for the reverse saturation current I_s will also increase. Eventually, their velocity and associated kinetic energy ($W_K = \frac{1}{2}mv^2$) will be sufficient to release additional carriers through collisions with otherwise stable atomic structures. That is, an *ionization* process will result whereby valence electrons absorb sufficient energy to leave the parent atom. These additional carriers can then aid the ionization process to the point where a high *avalanche* current is established and the *avalanche breakdown* region determined.

The avalanche region (V_Z) can be brought closer to the vertical axis by increasing the doping levels in the *p*- and *n*-type materials. However, as V_Z decreases to very low levels, such as -5 V, another mechanism, called *Zener breakdown,* will contribute to the sharp change in the characteristic. It occurs because there is a strong electric field in the region of the junction that can disrupt the bonding forces within the atom and "generate" carriers. Although the Zener breakdown mechanism is a significant contributor only at lower levels of V_Z, this sharp change in the characteristic at any level is called the *Zener region* and diodes employing this unique portion of the characteristic of a *p-n* junction are called *Zener diodes.* They are described in detail in Section 1.14.

The Zener region of the semiconductor diode described must be avoided if the response of a system is not to be completely altered by the sharp change in characteristics in this reverse-voltage region.

The maximum reverse-bias potential that can be applied before entering the Zener region is called the peak inverse voltage (referred to simply as the PIV rating) or the peak reverse voltage (denoted by PRV rating).

If an application requires a PIV rating greater than that of a single unit, a number of diodes of the same characteristics can be connected in series. Diodes are also connected in parallel to increase the current-carrying capacity.

Silicon versus Germanium

Silicon diodes have, in general, higher PIV and current rating and wider temperature ranges than germanium diodes. PIV ratings for silicon can be in the neighborhood of 1000 V, whereas the maximum value for germanium is closer to 400 V. Silicon can be used for applications in which the temperature may rise to about 200°C (400°F), whereas germanium has a much lower maximum rating (100°C). The disadvantage of silicon, however, as compared to germanium, as indicated in Fig. 1.23, is the higher

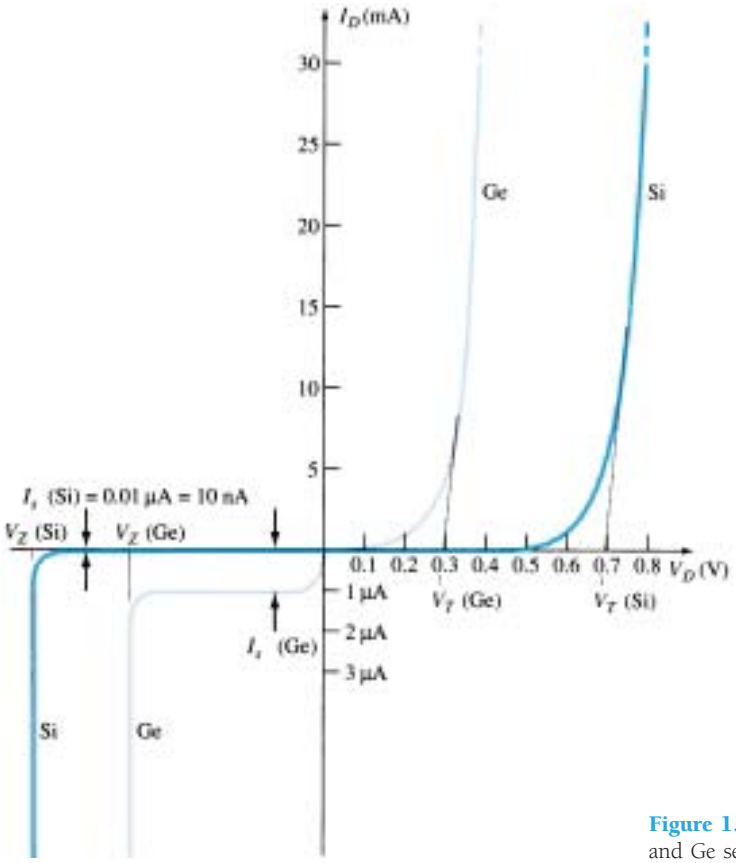


Figure 1.23 Comparison of Si and Ge semiconductor diodes.

forward-bias voltage required to reach the region of upward swing. It is typically of the order of magnitude of 0.7 V for *commercially* available silicon diodes and 0.3 V for germanium diodes when rounded off to the nearest tenths. The increased offset for silicon is due primarily to the factor η in Eq. (1.4). This factor plays a part in determining the shape of the curve only at very low current levels. Once the curve starts its vertical rise, the factor η drops to 1 (the continuous value for germanium). This is evidenced by the similarities in the curves once the offset potential is reached. The potential at which this rise occurs is commonly referred to as the *offset, threshold,* or *firing potential.* Frequently, the first letter of a term that describes a particular quantity is used in the notation for that quantity. However, to ensure a minimum of confusion with other terms, such as output voltage (V_o) and forward voltage (V_F), the notation V_T has been adopted for this book, from the word "threshold."

In review:

$$V_T = 0.7 \text{ (Si)}$$
$$V_T = 0.3 \text{ (Ge)}$$

Obviously, the closer the upward swing is to the vertical axis, the more "ideal" the device. However, the other characteristics of silicon as compared to germanium still make it the choice in the majority of commercially available units.

Temperature Effects

Temperature can have a marked effect on the characteristics of a silicon semiconductor diode as witnessed by a typical silicon diode in Fig. 1.24. It has been found experimentally that:

The reverse saturation current I_s will just about double in magnitude for every 10°C increase in temperature.