Chapter 1. ***PN Junction Diode Theory and Applications***

To understand the operation of a semiconductor (PN junction) diode, upon which BJT and MOS transistor operation depends, we must first understand the process by which electric current flows through solids. Different kinds of solids vary tremendously in the type and number of charge carriers that are available, and in the ease with which these carriers move when in the presence of an applied electric field. A solid may be classified into one of the following three groups:

A. ***Insulator*** - has almost NO available charge carriers.

B. ***Conductor*** - has large numbers of mobile charge carriers that can easily be moved in the presence of an electric field.

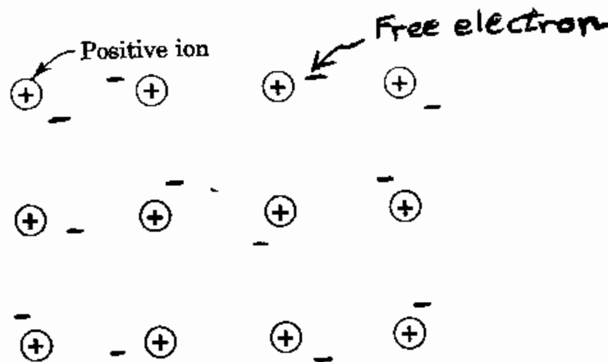
C. ***Semiconductor*** - has conduction properties that are between those of an insulator and a conductor.

1.1 Metallic Conductors

In a typical metal, such as copper or silver, the atoms are arranged in a systematic array to form a *crystal*. The atoms are packed in such close proximity (compared to the distance from their outer shell electrons to their nucleus) that the outer shell, loosely-bound electrons are attracted with equal strength to the neighboring nuclei. Therefore they cannot be clearly associated with any one metal atom's nucleus.

Thus these outer-shell, or ***conduction*** electrons, are free to wander through the crystal's "lattice" structure shown in Fig. 1-1.

Fig 1-1. Simplified 2-dimensional model of a metallic crystal.



At *absolute zero* temperature (0 Kelvin degrees) the conduction electrons experience no opposition to their motion, and thus the resistance of the conductor is zero.

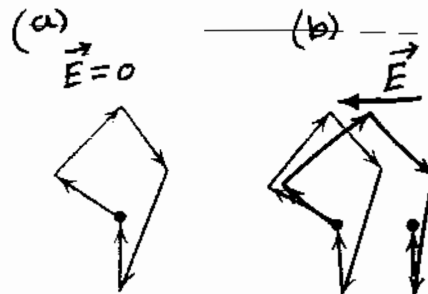
However, at temperatures above 0 K, the atoms (minus their wandering outer shell electrons) possess kinetic (thermal) energy in the form of vibration about their neutral positions in the lattice. The higher the temperature, the larger will be the vibrational displacement about the neutral position. Because each of these bound atoms lack their outer-shell electrons, they are called *positive bound ions*, since they must have a net positive charge.

Even with no applied electric field, the conduction electrons will collide with the thermally vibrating bound positive ions, and thus will execute a "random walk", where there is no net motion, and thus the net current is zero, as shown in Fig 1-2a. But if a uniform electric field intensity E (units of V/m) is applied across the conductor, then a small *drift velocity* "u" in the direction of -E is superimposed upon the random motion. This is because the electric field exerts a force on a charged particle according to

$$F = q \cdot E$$

where q is the charge on an electron, which is negative (q = -1.6 X 10^{-19} Coulombs). At each inelastic collision with an ion, the electron loses its kinetic energy. It then accelerates again, gaining a component of velocity in the -E direction, and then it looses its energy (and velocity) at the next inelastic collision, as shown in Fig. 1-2b.

Fig. 1-2 Motion of an electron in a metallic crystal (a) without E field (b) with E field



The average time between collisions, ***Tc***, is determined solely by the random velocity and the average length of the free path between collisions with lattice ions (mean free path). Because the average random velocity is constant at a given temperature, the mean free path depends only upon the spacing of lattice ions in the crystal, and thus is a material-dependent constant. Specifically note that ***Tc*** does NOT depend upon the strength of the applied electric field, ***E***, since the conduction electron's random velocity component is ***much greater*** than its field-induced (drift) velocity component. Because the average field-induced velocity component varies linearly from 0 (just after a collision) to ***a*Tc*** (just before the next collision); and because between collisions the acceleration of the electron is governed by Newton's law of force, ***F = ma = qE***, it is clear that the average drift velocity must be given by ***a*(Tc)/2*** where the acceleration ***a = qE/m***.

Therefore the conduction electron maintains an average drift velocity given by ***u = qE(Tc)/(2m)***, that is directed along ***-E***. Clearly the drift velocity u is directly proportional to the magnitude of ***E***. The drift velocity is given by

$$u = \frac{q \cdot Tc}{2 \cdot m} \cdot E = \mu \cdot E \qquad \text{Where} \quad \mu = \frac{q \cdot Tc}{2 \cdot m} \qquad (1-1)$$

The constant of proportionality, μ, is called the ***mobility***. It has units of ***(m/s)/(v/m) = m²/V-s***. *It dep*ends upon the material and the temperature.

The resulting flow of electrons, each carrying charge "***q***" (in Coulombs) with drift velocity ***u*** (in ***m/s***) constitute an electric current. If there are ***n*** free (conduction) electrons per cubic meter, then the current density ***J*** (in ***A/m²***) is given by

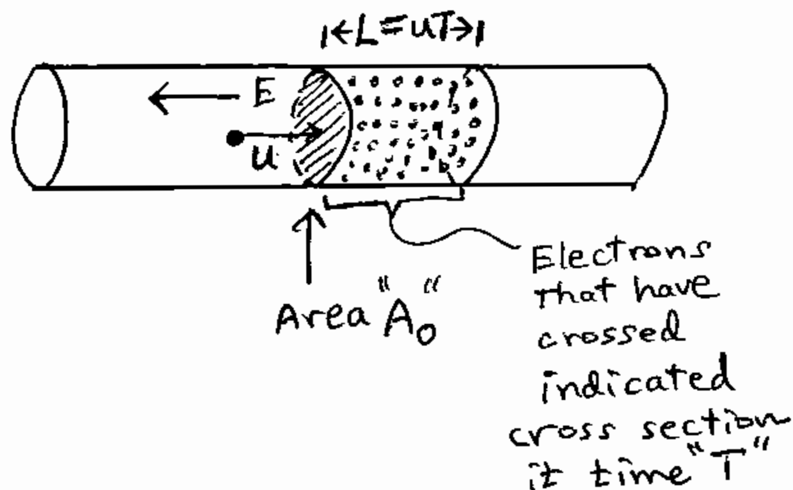
$$J = n \cdot q \cdot u = n \cdot q \cdot (\mu \cdot E) = \sigma \cdot E \qquad \text{Where} \quad \sigma = n \cdot q \cdot \mu \qquad (1-2)$$

The constant of proportionality σ is the *conductivity* of the material in units of ***(A/m²)/(V/m) = (A/V)/m = S/m***.

****

Proof: Consider the cylindrical conductor of uniform cross-section shown in Fig. 1-3. Note the cross section labeled "area A_0". If an electric field is directed along the axis of the conductor, then all free electrons will move with average drift velocity ***u*** to the right. We can think of this as a pipe conveying a moving "electron gas" to the right. Then, in a T-second time interval, the volume of this "electron gas" that crosses the indicated cross section is A_0L, where ***L = uT***.

Fig. 1-3 Current density calculated using "moving electron gas" model



The total charge ***Q*** that crosses the indicated cross section in ***T*** seconds is found by multiplying by the electron density, ***n*** (in units of $1/m^3$) times the charge on one electron, ***q*** (in units of Coulombs), hence

$$Q = A_0 \cdot u \cdot T \cdot n \cdot q \quad \text{and thus} \quad J = \frac{\frac{Q}{T}}{A_0} = n \cdot q \cdot u \quad \left(\text{units of } \frac{A}{m^2}\right)$$

(Handwritten annotations: "Volume" over $A_0 \cdot u \cdot T$, "L" over $u \cdot T$, "electron density" pointing to n, "charge per electron" pointing to q.)

****

Solving (1-2) for E yields

$$E = \frac{1}{\sigma} \cdot J = \rho \cdot J \tag{1 - 3}$$

where $\rho \triangleq \frac{1}{\sigma}$ and is called the ***resistivity***

of the material. It has the units of ***(V/m)/(A/m²) = Ω-m***

Equations (1-2) and (1-3) are a microscopic statement of ***Ohm's Law***. For a conductor of uniform cross section ***A (m²)*** and length ***L (m)***,

$$I = J \cdot A = \sigma \cdot A \cdot E = \sigma \cdot A \cdot \frac{V}{L} = \sigma \cdot \frac{A}{L} \cdot V = G \cdot V \quad \left(\text{where } G \triangleq \frac{\sigma A}{L}\right) \tag{1 - 4}$$

Where ***G*** is the ***conductance*** of the conductor in ***Siemens*** (or ***mhos***).

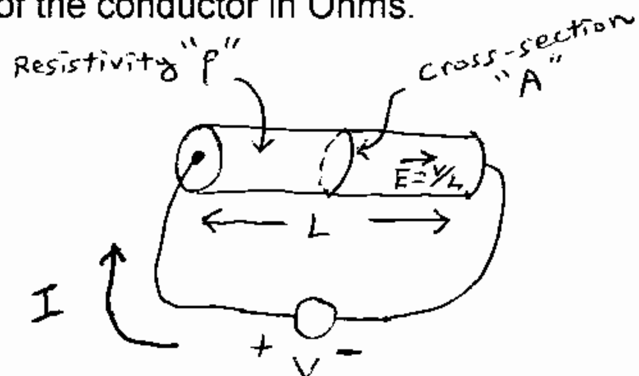
Solving (1-4) for V, we find that

$$V = \frac{1}{\sigma} \cdot \frac{L}{A} \cdot I = \frac{\rho \cdot L}{A} \cdot I = R \cdot I \quad \left(\text{Macroscopic Form of Ohm's Law}\right) \tag{1 - 5}$$

Where ***R*** is the ***resistance*** of the conductor in Ohms.

Note:

$$R = \frac{\rho L}{A}$$



****

Example 1-1. The resistivity of copper at 20 deg. C is

$$\rho = 1.73 \cdot 10^{-8}\ \Omega \cdot m$$

Find the average drift velocity in a copper conductor with a cross-sectional area of 10^{-6} m^2 that is carrying a current of 10 A.

Solution: Cu has an atomic weight of 63.6 and a density of 8.9 g/cm³. Therefore, by ***Avogadro's Law***, the number of atoms per meter is

63.6 g of Cu has Avagadro's number of atoms in it.

$$n_A = \frac{6.022 \cdot 10^{23} \cdot \quad 8.9 \cdot 10^{6} \cdot \frac{grams}{m^3}}{63.6\ grams}$$

$$n_A = 8.427013 \cdot 10^{28} \cdot m^{-3}$$

# of Cu atoms / cubic meter of Cu

Assuming that there is one free electron per atom, then the number of free electrons per cubic meter is

$$n = n_A$$

Therefore by (1-2),

$$q = 1.6 \cdot 10^{-19} \quad \frac{\text{Coulombs}}{\text{electron}}$$

$$u = \frac{J}{n \cdot q} = \frac{\frac{I}{A}}{n \cdot q} = \frac{\frac{10}{10^{-6}}}{8.43 \cdot 10^{28} \cdot 1.6 \cdot 10^{-19}} = 7.41 \cdot 10^{-4}\ \frac{m}{s}$$

****

Thus the average drift velocity in a good conductor is ***very slow*** *com*pared to the random thermal electron velocities, which are on the order of ***10^5 m/s*** at room temperature.

Finally, note that as temperature increases, random thermal motion increases, and therefore the average time between collisions, ***Tc***, decreases. Therefore as temperature increases, by (1-1) the mobility decreases; and by (1-2) the conductivity therefore decreases.

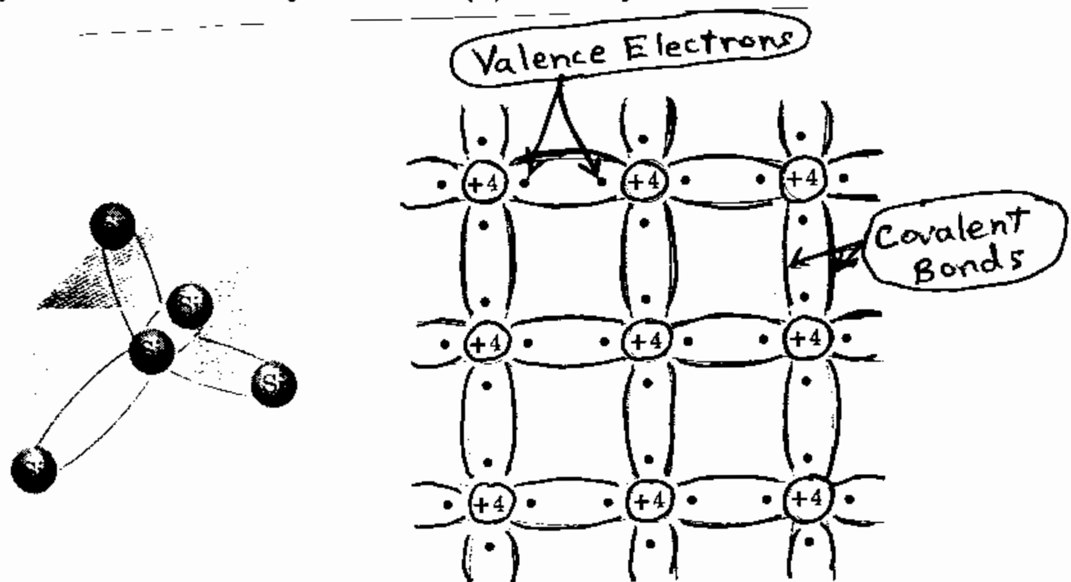
Recall $R = \frac{\rho L}{A}$

GENERAL CONCLUSION: ***the resistance of metallic conductors increases with increasing temperature.***

1.2 Intrinsic Semiconductors

The semiconductor of greatest importance in electronics ***silicon*** (Si). It is located in the fourth column of the ***periodic table***, and thus has four ***valence*** (***conduction***) electrons in its outer shell. The crystal structure of Si follows a ***tetrahedral*** pattern, with each atom sharing one valence electron with each of the four neighboring atoms. The covalent bonds are shown in the two-dimensional representation of Fig. 1-4.

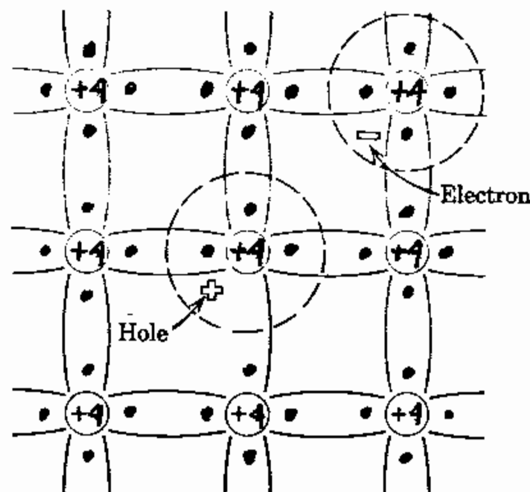
Fig. 1-4. Atomic structure of a silicon crystal at 0K
(a) 3-D tetrahedral pattern (b) 2-D representation



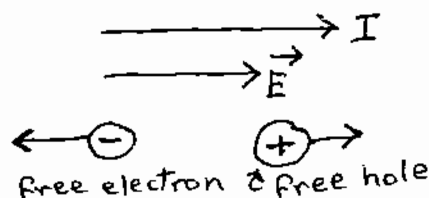
At temperatures close to absolute zero, the electrons in the outer shell are tightly bound. Thus there are no free charge carriers, and silicon is an insulator.

The energy required to break a covalent bond of silicon is about ***1.1 electron volts*** (***eV***), where 1 eV is the amount of kinetic energy imparted to a single electron when it is accelerated through a 1-volt potential difference. At room temperature (300 K), a few electrons have this amount of thermal energy, and thus are "excited into the conduction band", that is, they are "broken loose by the thermal motion" from their covalent bond, thereby becoming a free (conduction) electron. When a covalent bond is broken (see Fig. 1-5), a vacancy, or ***hole*** is left behind.

Fig. 1-5 Silicon Crystal at room temperature showing a covalent bond broken: a free "electron-hole" pair is generated.



A valence electron from a neighboring covalent bond can easily fill the hole (without ever gaining sufficient energy to become "free"). Thus the electron vacancy appears in a new place, and the effect is as if a positive charge of magnitude q has moved to a new location. Thus both the negatively charged electron and the positively charged hole are mobile, and can move under the influence of an applied electric field.



***General Conclusion**: Conduction in semiconductors is due to two separate and independent "particles", **holes** and **electrons**, which carry opposite charges of equal magnitude (**q**) and drift in opposite directions under the influence of an applied electric field.*

With two charge-carrying particles, the expression of the current current density becomes:

$$J = \left(n \cdot \mu_n + p \cdot \mu_p\right) \cdot q \cdot E = \sigma \cdot E \tag{1 - 6}$$

where ***n*** and ***p*** are the concentrations of electrons and holes (***number/m³)***, and μ_n and μ_p are the mobilities of the electrons and holes (***m²/(V-s)***). Note that the conductivity σ is

$$\sigma \triangleq \left(n \cdot \mu_n + p \cdot \mu_p\right) \cdot q \tag{1 - 7}$$

Electric Charge Neutrality Condition

Because holes and electrons are created simultaneously, in a pure (intrinsic) semiconductor the ***total number of holes (mobile positive charges) must equal the total number of electrons (mobile negative charges)***, thus we require. Thus we may state that in a pure semiconductor

$$n = p = n_i \tag{1 - 8}$$

Where ni is the "***intrinsic carrier concentration***"

New electron-hole pairs are created continuously. Under equilibrium conditions, they are removed by recombination at the same rate. At room temperatures, only a very small fraction of the valence electrons are in the conduction state at any instant. We see from Table 1-1 that n_i is very small compared to the total number of Cu atoms/m³.

Table 1-1 Some useful constants

Avogadro's Number	$AvNr := 6.022 \cdot 10^{23} \frac{atoms}{g \cdot atom}$
Charge on an electron	$q := 1.6 \cdot 10^{-19}\ C$
Mass of an electron	$m := 9.109 \cdot 10^{-31}\ kg$
Energy Gap of Si at 300K	$E_g = qVg := 1.12\ eV$
Electron Mobility of Si at 300K	$\mu_n := 0.135 \frac{m^2}{V \cdot s}$
Hole Mobility of Si at 300K	$\mu_p := 0.048 \frac{m^2}{V \cdot s}$
Intrinsic Carrier Density of Si at 300K	$n_i := 1.5 \cdot 10^{16}\ m^{-3}$
Intrinsic Resistivity of S at 300K	$\rho_i := 2300\ \Omega \cdot m$
Density of Si at 300K	$dens_Si := 2.33 \cdot 10^6\ g \cdot m^{-3}$
Atomic Weight Si	28.09

****

Example 1-2 Find the relative concentration of Si atoms and electron-hole pairs at room temperature. Then estimate the intrinsic resistivity of Si.

Solution: By Avogadro's Law, the number of atoms, n_A, in one cubic meter of Si is

$$n_A = \frac{6.022 \cdot 10^{23} \cdot 2.33 \cdot 10^6}{28.09} \qquad \left(\frac{Si\ atoms}{g}\right) \left(g/m^3\right)$$

$$n_A = 4.995109 \cdot 10^{28} \frac{Si_Atoms}{m^3}$$

Room Temperature in Kelvin $T := 300 \cdot K$

Bandgap of Silicon: $Eg := 1.12 \cdot eV$ (From SS Text)

Boltzmann's Constant: $k := 8.62 \cdot 10^{-5} \cdot \frac{eV}{K}$ (From SS Text)

For Silicon $B := 5.4 \cdot 10^{31} \cdot cm^{-6} \cdot K^{-3}$

$$n_i := \left(B \cdot T^3 \cdot \exp\left(\frac{-Eg}{k \cdot T}\right)\right)^{\frac{1}{2}}$$

Intrinsic Concentration Formula From Equation 3.2 (SS Text)

$$n_i := 1.5 \cdot 10^{10} \cdot \frac{1}{cm^3}$$

Or $n_i = 1.5 \times 10^{16} \frac{1}{m^3}$

Thus $\frac{n_i}{n_A} = 3.002938 \cdot 10^{-13}$ electron-hole pairs/Si atoms (at room temperature)

Since in pure Si (1-8) tells us , $n = p = n_i$

Then (1-7) tells us that the intrinsic conductivity is

$$\sigma_i := n_i \cdot (\mu_n + \mu_p) \cdot q$$

$$\sigma_i = 4.392 \cdot 10^{-4} \quad \frac{S}{m}$$

Then the intrinsic resistivity is $\rho_i := \frac{1}{\sigma_i}$

$$\rho_i = 2.276867 \cdot 10^{3} \quad \Omega \cdot m$$

Which is close to the given value for ρ_i in Table 1-1.

****

The above example confirms that intrinsic Si is a poor conductor, as compared to copper (σ = 5.8 X 10^7 S/m). Why then is Si of such interest in electronics? The answer lies in the fact that Si exhibits two unusual and useful electrical properties:

1. The concentration of free carriers, and consequently the conductivity *increases exponentially with temperature* (about 5% per degree C at room temperatures). The *thermistor* is a temperature-sensitive device that makes use this property to serve as a very sensitive temperature measuring device for use in instrumentation and control applications. It also is used to compensate for the decrease in conductivity with increasing temperature of metallic parts of a circuit, such as in protecting cold CRT filaments from large initial "inrush" currents (and possible burnout) upon power-up.

In this application, the thermistor, which is called a "Globar resistor", is placed in series with the CRT filament. When the power is turned on, the cold filament has a low resistance, but the cold thermistor has a considerably higher resistance, thus the initial current is kept low. At this point, the filament does not glow very brightly, because the thermistor has the higher resistance, and thus dissipates i^2R heat, since the current is the same through both the filament and the thermistor. But as both the thermistor and filament warm up, due to the current flow through them, the situation is reversed; the hot thermistor now exhibits a resistance that is much lower than the resistance of the hot filament, and the situation is the same as if the thermistor were not in the circuit.

2. The conductivity of a semiconductor can be increased greatly, and precisely controlled, by adding small amounts of certain impurities in a process called *doping*. There are two types of ***doping impurities***, ***n-type*** and ***p-type***. The first greatly increases the population of free electrons, while the second greatly increases the population of mobile holes. Doping will be discussed in the next section.

1.3 Doped Semiconductors

Doping is the process by which impurity atoms from either the third column (elements with 3 outer-shell, or valence, electrons) or the fifth column (elements with 5 valence electrons) of the Periodic Table are introduced into the Si crystal lattice, which is found in the fourth column (elements with 4 valence electrons) of the Periodic Table (see Fig. 1-5.)

1.3.1 n-type Silicon

If a small amount of ***pentavalent impurity*** element (***antimony***, ***phosphorus***, or ***arsenic***) is introduced into an otherwise pure silicon crystal lattice, only four of the outer five electrons of the impurity atom are needed to complete the covalent bonds.

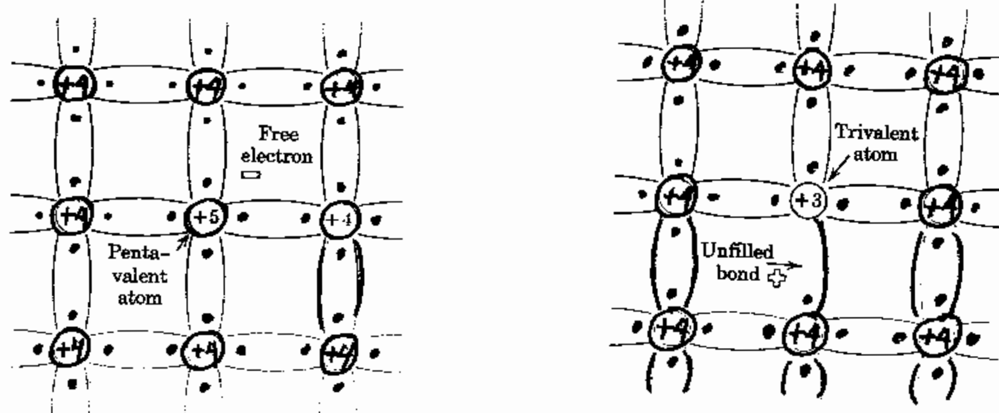
Fig. 1-5. Semiconductor Elements in the Periodic Table

acceptor impurity → III; donor impurity → V

III (+3)	IV (+4)	V (+5)
5 B BORON 10.82	6 C CARBON 12.01	7 N NITROGEN 14.008
13 Al ALUMINUM 26.97	14 Si SILICON 28.09	15 P PHOSPHORUS 31.02
31 Ga GALLIUM 69.72	32 Ge GERMANIUM 72.60	33 As ARSENIC 74.91
49 In INDIUM 114.8	50 Sn TIN 118.7	51 Sb ANTIMONY 121.8

Therefore, the remaining electron, along with the addition of a small amount of thermal energy, becomes a free electron available for conduction. The resulting material is called an ***n-type semiconductor***, because the majority of the available charge carriers are now electrons, with a negative charge. Of course, the entire crystal is still electrically neutral, since for every conduction electron (charge -q) that moves away from its impurity site, a positive ion is left behind (charge +q) that is bound to the impurity site. Thus in n-type semiconductor, the positive ions are ***immobile***, and cannot contribute to current flow in the doped semiconductor. Fig. 1-6a shows the effect of n-type doping. A pentavalent impurity atom is called a ***donor atom***, because it *donates* an electron to the lattice. With a small amount of energy, this donated electron can be conducted away from the impurity site, thereby "***uncovering***" a bound positive lattice ion.

Fig. 1-6. Doped Silicon Semiconductor
(a) Effect of n-type doping. (b) Effect of p-type doping



1.3.2 p-type Silicon

If a small amount of ***trivalent*** impurity (aluminum, boron, gallium, or indium) is introduced into an otherwise pure Si crystal lattice, a ***p-type semiconductor*** results. As shown in Fig. 1-6b, when a trivalent atom replaces a silicon atom in the lattice, only three valence electrons are available to complete the covalent bonds. If the remaining unfilled covalent bond is filled by a bound valence electron from a neighboring atom, a positively-charged mobile ***hole*** is created that can effectively travel from atom to atom under the influence of an external electric field. The trivalent atom is called an ***acceptor atom***, because it *accepts* an electron as the hole "moves away" from the impurity site. As the hole moves away, it leaves behind, or "***uncovers***", an ***immobile*** negative ion that is **bound** to the impurity site.

1.3.3 Conduction in a Doped Semiconductor

By adding donor or acceptor atoms in small amounts, the conductivity of a semiconductor can be increased dramatically. In order to be able to calculate the conductivity of a doped semiconductor, we must first understand the processes of ***thermal generation*** and ***recombination***.

Thermal Generation Rate

Depends upon the properties of the material and its temperature. It is denoted by the symbol "***g***" and has units of "***electron-hole pairs per second per cubic meter***".

The energy required for a valence electron to become a free electron, or the energy needed in the "generation" of a conduction electron can be expressed in electron volts as ***qVg***, where ***Vg*** is the potential difference between the conduction band and the valence band. The average energy that corresponds to a temperature ***T*** can be expressed as ***kT***, where ***k*** is ***Boltzmann's constant***

$$k = 8.620 \cdot 10^{-5}\ \frac{eV}{K} \quad \text{or} \quad 1.381 \cdot 10^{-23}\ \frac{J}{K}$$

Statistical thermodynamics shows that the probability of a valence electron receiving enough thermal energy to become free is proportional to

$e = q = 1.6 \times 10^{-19}$ C

$E_g = V_g$ = Bandgap Energy (= 1.1 eV for Si)

$$e^{-\frac{q \cdot V_g}{k \cdot T}} \qquad (1 - 9)$$

Hence the rate of thermal generation is proportional to this term, and is thus a strong function of temperature, T..

If a hole collides with a mobile electron, they will ***recombine*** and thus both the hole and the mobile electron will disappear. The electron may be thought of as "falling into" the hole. If there are only a few electron-hole pairs in existence, the rate of recombination is low, and if there are many, the rate of recombination is high. If, as in n-type material, there are few holes but many mobile electrons, the rate of recombination will still be high, due to the large number of electrons. In general the rate of recombination R is given by the ***mass-action law***:

$$R = r \cdot n \cdot p \qquad (1 - 10)$$

where ***r*** is a proportionality constant for the material, and ***n*** is the number of mobile electrons per cubic meter and ***p*** is the number of holes per cubic meter.

Under equilibrium conditions, the rate of generation just equals the rate of recombination

$$g = R = r \cdot n \cdot p \qquad (1 - 11)$$

In a pure Si crystal, the intrinsic concentrations of electrons and holes are equal, hence

$$g = R = r \cdot n_i \cdot p_i = r \cdot n_i^2 \qquad (1 - 12)$$

Even in a doped semiconductor crystal, the great majority of the atoms are still silicon; only one in perhaps ten million atoms are dopant (impurity) atoms. Thus the thermal generation rate is essentially unchanged from its intrinsic value. Thus (1-12) still predicts the thermal generation rate in a doped semiconductor. However now the mobile electron and hole concentrations, ***n*** and ***p***, can be quite different. Thus we may write

$$g = r \cdot n \cdot p = r \cdot n_i^2 \quad \Rightarrow \quad n \cdot p = n_i^2 \qquad (1 - 13)$$

Equation (1-13) shows that in any doped semiconductor, the product of the free electron density (n) and the hole density (p) is a constant that remains equal to the electron-hole pair density that would exist in an undoped (intrinsic) sample of that same semiconductor at the same temperature.

Furthermore, recall that the total crystal must remain electrically neutral. There are two types of charge to consider:

(1) the immobile lattice ions
(2) the mobile charge carriers.

At normal device operating temperatures, practically all donor and acceptor atoms are ionized (uncovered), leaving the immobile positive and negative ions in concentrations equal to the original doping levels, N_d and N_a, respectively.

Also, the mobile positive and negative mobile charge carriers have the concentrations p and n, respectively. Then the electrical charge neutrality requirement dictates that in any doped crystal Total(+) = Total(−)

$$p + N_d = n + N_a \qquad (1-14)$$

In n-type material, which is created by adding donor impurity atoms to intrinsic semiconductor, the acceptor impurity concentration is zero. Hence Na = 0, and (1-14) becomes

$$n_n = p_n + N_d \qquad (1-15)$$

Note that the "n" subscripts were added to emphasize that (1-15) pertains to n-type semiconductor material.

In practical cases, the doping level is made large enough that Nd >> p_n, and thus (1-15) may be approximated by

$$n_n \cong N_d \qquad (1-16a)$$

$n_n p_n = n_i^2$

But then (1-13) implies that

$$p_n = \frac{n_i^2}{n_n} \cong \frac{n_i^2}{N_d} \qquad (1-16b)$$

Using a similar line of reasoning for p-type material, it is found that

$$p_p \cong N_a \quad \text{and} \quad n_p \cong \frac{n_i^2}{N_a} \qquad (1-17)$$

Example 1-3 **

Investigate the properties of a silicon crystal if every 10 millionth silicon atom has been replaced by an atom of indium.

Solution: in Example 1-2 we found that the number of silicon atoms per cubic meter is

$$n_A = 4.995109 \cdot 10^{28} \quad \frac{\text{SiAtoms}}{m^3}$$

Indium is a trivalent atom, therefore it is an acceptor impurity, which results in a p-type semiconductor. The required indium doping concentration is

$$N_a = n_A \cdot 10^{-7}$$

$$N_a = 4.995109 \cdot 10^{21} \quad \frac{\text{Indium atoms}}{m^3}$$

But from (1-17), this doping level forces the hole concentration to approximately equal the value of $\boldsymbol{N_a}$

$$p_p \cong N_a$$

$$p_p \cong 5 \cdot 10^{21} \ \frac{\text{holes}}{m^3}$$

Then (1-17) predicts that the free electron concentration is

$$n_p \cong \frac{n_i^2}{p_p}$$

$$n_p \cong 4.5 \cdot 10^{10} \ \frac{\text{electrons}}{m^3}$$

The conductivity of the doped Si is found from (1-7)

$$\sigma = \left(n_p \cdot \mu_n + p_p \cdot \mu_p\right) \cdot q$$

$$\sigma = 38.4 \quad \frac{S}{m}$$

The resistivity is

$$\rho = \frac{1}{\sigma}$$

$$\rho = 0.026042 \quad \Omega \cdot m$$

Note how this compares with the resistivity of intrinsic silicon (found in Table 1-1), ρ_i = 2300 Ω-m.

It is truly remarkable that such an extremely minute doping level (1 part in 10 million) can result in such a profound change (lowering) in resistivity!

****

1.3.4 Diffusion in a Doped Semiconductor

If the doping is nonuniform, the concentration of charged particles will also be nonuniform, and charge motion will result as a result of the thermally-driven statistical "***random-walk***" process called ***diffusion***. We are all familiar with this process by the way the smell of perfume can diffuse throughout a room when a perfume bottle is opened at one end of a room. For this process to occur, two contributing factors must be present:

(1) a concentration gradient

(2) random thermal motion

Due to the random thermal motion, which is equally likely to move each of the perfume molecules in ***any*** direction, there will be a *statistical net mass transfer from a region of higher perfume molecule concentration toward a region of lower concentration.* Imagine a room with an imaginary plane dividing its right and left halves. Then if there are more perfume molecules in the right half of the room, and fewer in the left half, then on a statistical basis, more molecules will cross from the right to the left half of the room (since more are available to do this) than will cross from the left to the right half of the room.

If the diffusion process is allowed to continue, unopposed, eventually the perfume molecules will be equally distributed around the room, and there will no longer be any concentration gradient.

Therefore, if a nonuniform concentration of electrons and holes undergoing random thermal motion exists, then diffusion will occur.

Considering first just the electron concentration gradient, there will be a net movement of electrons from the denser region to the other, resulting in an ***electron diffusion current*** that is proportional to the concentration gradient dn/dx, where x is the direction along which the concentration is changing. The electron diffusion current density (A/m²) is given by

$$J_n = q \cdot D_n \cdot \frac{d}{dx} n \qquad (1 - 18a)$$

where D_n is the ***diffusion constant*** for electrons (m²/s), which depends upon the semiconductor material and its temperature. If dn/dx is positive, the resulting electron motion will be in the -x direction, thus constituting an electrical current in the +x direction.

The corresponding equation for hole diffusion current density is

$$J_p = -q \cdot D_p \cdot \frac{d}{dx} p \qquad (1 - 18b)$$

The minus sign is needed because the flow of positive charged holes is in the direction of the ***negative*** concentration gradient. '

Note that in both cases, the charged particles diffuse ***away from*** the region of greatest concentration - but the motion is ***NOT*** due to any force of repulsion - it is due purely to the results of a statistical random walk.

It makes sense that the carrier diffusion constants should depend upon the average thermal energy level and also the carrier mobilities. Therefore, it should come as no surprise that the diffusion constant (D) is the product of carrier mobility (μ) and the average thermal energy per coloumb of charge that is carried (kT/q). Thus, we require that

$$D_n = \mu_n \cdot \frac{k \cdot T}{q} \qquad D_p = \mu_p \cdot \frac{k \cdot T}{q} \qquad (1-19)$$

These two equations may be combined to form the ***Einstein Relation***:

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{k \cdot T}{q} \qquad (1-20)$$

1.4 PN Junction Diode

In a pure or doped semiconductor (p-type or n-type), current flows with equal ease in either direction. Current flow through a bulk semiconductor (or conductor) is said to be ***bilateral***.

However, if a p-type semiconductor is suddenly butted up against an n-type semiconductor, a ***pn junction semiconductor diode*** is formed, and there exists a carrier density gradient that is ***unilateral***; and current can only flow easily in one direction.

1.4.1 pn Junction Diode Fabrication

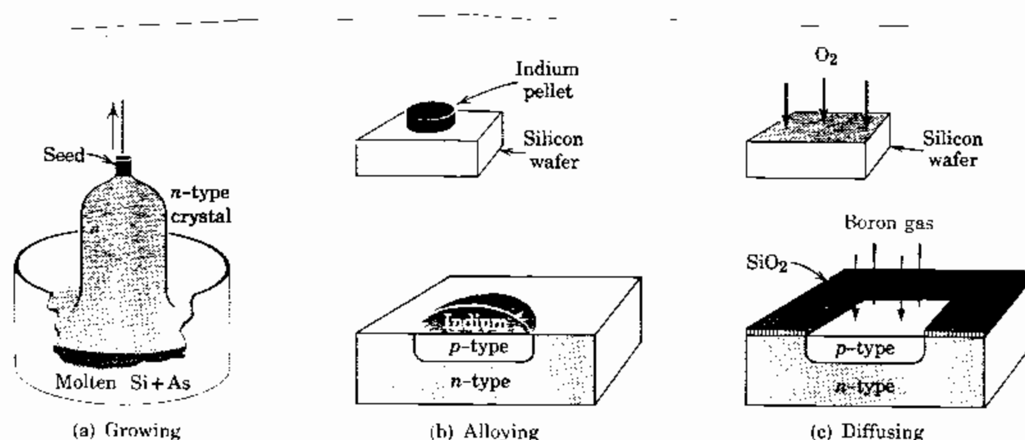
A pn junction diode can be fabricated by either ***alloying*** or by ***gas diffusion***, as shown in Fig. 1-7. First a doped silicon crystal is grown. This is done (Fig. 1-7a) by melting silicon with a very small amount of dopant (a pentavalent impurity such as arsenic is used to make n-type silicon). A "seed" crystal is touched to the surface of the molten silicon, then it is slowly raised.

The n-type silicon crystallizes onto the seed, and it grows in the same orientation as the seed crystal. The resulting n-doped crystal is called a ***boule***, and it resembles a slab of salami, perhaps ten centimeters in diameter. This slab is then sliced (using a diamond saw) into thin wafers of n-type Si.

If a dot of trivalent impurity such as indium is placed on the wafer and then heated to the proper temperature, and a small drop of molten material will form at the top surface of the wafer (Fig. 1-7b). Upon cooling the alloy solidifies (crystallizes) into p-type silicon (since there are many more acceptor impurities than donor impurities in this portion of the silicon). Thus an ***alloyed pn junction diode*** results.

An integrated circuit diode is formed using the principle of gas diffusion (Fig. 1-7c). First, the n-type silicon wafer is heated in the presence of oxygen to grow a thin layer of insulating (glassy) ***silicon dioxide*** (***$Si0_2$***) on the surface of the wafer. Then a selected portion of the oxide layer is removed by etching through a photographically defined mask. When the heated wafer is exposed to boron gas in a diffusion furnace, the boron acceptor atoms penetrate into the selected area of the wafer by solid-state diffusion, and a precisely-defined pn junction is formed. (The unetched oxide layer is impervious to boron gas.)

Fig. 1-7 Methods of forming pn junctions.



1.4.2 pn Junction Diode External Behavior

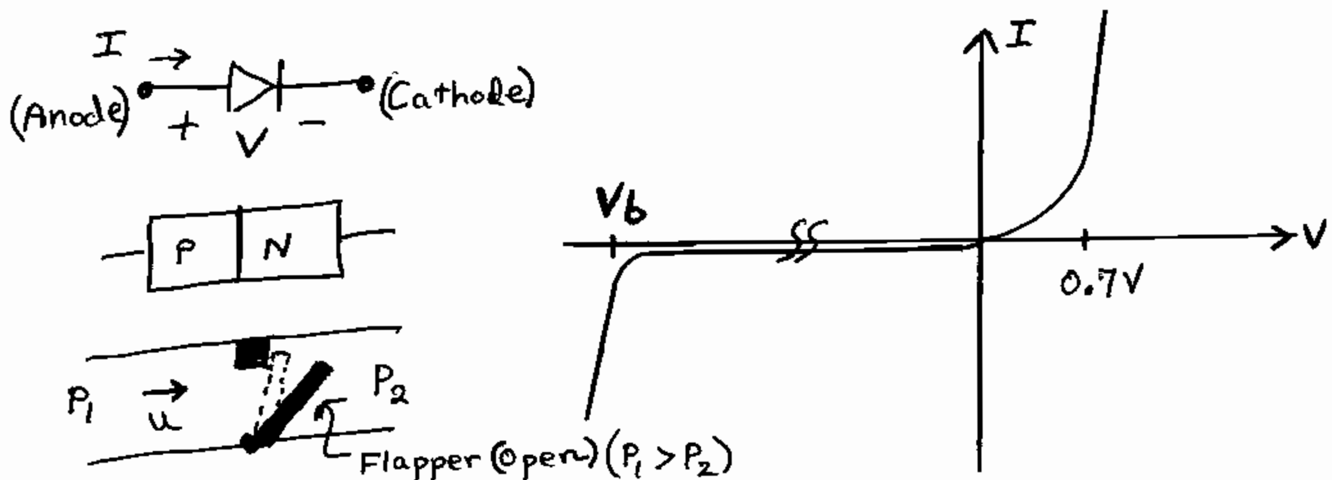
The external behavior of a junction diode is described in Fig. 1-8, which shows the I-V curve for the diode along with its schematic symbol and "***check valve***" hydraulic analogy. The arrow side (p-side) of the diode is called the ***anode,*** and the bar side (n-side) is called the ***cathode***.

When the diode voltage ***V*** is made more positive than about 0.6 V, (which sets up an "electrical pressure" that attempts to push positive charge from the ***p*** to the ***n*** region), current flows easily in the direction of the symbol arrow, and the diode is said to be "***forward-biased***". In the check valve analogy, under conditions of forward bias, the flapper of the check valve is pushed open. According to the diode's I-V curve, the diode acts like a 0.6 V dc battery with a fairly low internal resistance (since the I-V curve is so steep in this region).

When the diode voltage ***V*** is made negative (which sets up an electrical pressure that attempts to push positive charge from ***n*** to ***p*** regions) , the diode is said to be "***reverse-biased***", and very little current can flow, even as the diode voltage is made more and more negative. In the check valve analogy, under conditions of reverse bias, the flapper is pushed shut, and no water can flow through it, even for fairly large reverse pressures. According to the diode's I-V curve, only a very small constant "***reverse saturation current***" will flow for reasonably large negative terminal voltages. The reverse-biased diode exhibits a very high resistance, since the I-V curve is nearly horizontal.

If the diode voltage becomes too negative, however, the diode junction suffers "***reverse (zener or avalanche) breakdown***" at V = -Vb, and all of a sudden the diode begins to act like a -Vb volt battery with fairly low internal resistance, due to the steepness of the I-V curve in this region..

Fig. 1-8 Junction Diode I-V Curve, Symbol, Check Valve Analogy



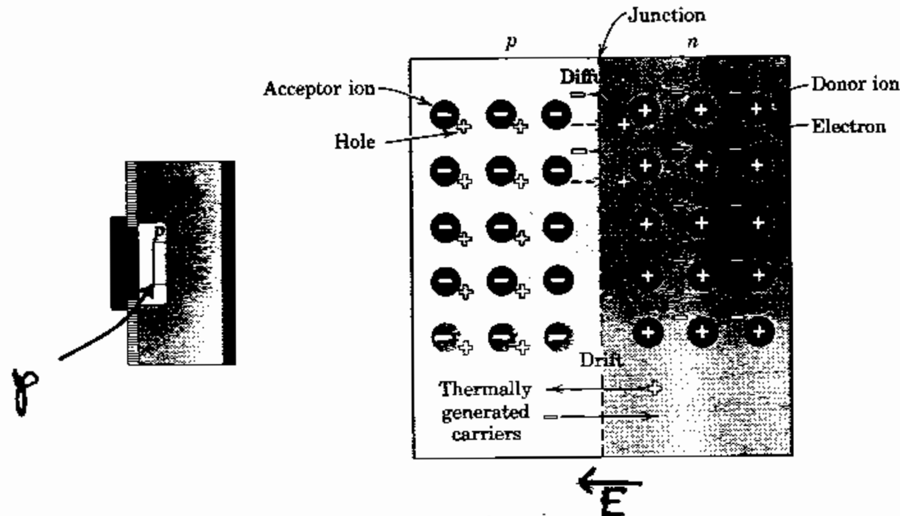
1.4.3 Internal Operation of the pn Junction

The distribution of charges in the vicinity of an open-circuited pn junction is depicted in Fig. 1-9. Note that the junction is located in the x = 0 plane. There will be no variation in the y and z directions. The circles represent the bound lattice ions: immobile positive donor ions in the n-type material and immobile negative acceptor ions in the p-type material. The separate + and - signs represent the mobile holes and electrons.

Imagine that the junction was just instantaneously created. Then the holes (mobile + charges) in the p-type semiconductor will begin to diffuse from right to left, across the junction into the n region, toward the direction of lower hole concentration. Once a hole crosses the junction, it soon recombines with one of the many free electrons in the n-type semiconductor.

In similar fashion, the free electrons in the n region will diffuse from left to right, across the junction into the p region, toward the direction of lower free electron concentration. Once a free electron crosses the junction, it soon recombines with one of the many free holes in the p-type semiconductor.

Fig. 1-9. Charges and conduction mechanisms in an open-circuited pn junction diode



The diffusion of holes from left to right uncovers bound negative acceptor ions to the left of the junction, and the diffusion of electrons from right to left ***uncovers*** bound positive donor ions to the right of the junction.

To simplify the remaining discussion, let us define holes in the p-region and electrons in the n-region to be ***majority carriers***, since they are in the majority. Likewise, holes in that have diffused into the n-region and electrons that have diffused into the p-region will be called ***minority carriers***, since they are in the minority in these regions.

The result of the diffusion of majority carriers is the uncovering of bound charge in the ***depletion region,*** which extends to either side of the original junction boundary, or ***metallurgical junction.*** This region of uncovered, bound charges of opposite signs on either side of the metallurgical junction sets up an internal electric field that is directed from right to left across the junction. This electric field causes minority carriers to drift back across the junction. Thus we have two oppositely-directed conduction mechanisms: drift and diffusion.

Initially, at the instant the junction is created, only diffusion current is present. But as more and more bound lattice ions are uncovered by diffusion, the internal electric field begins to build up, causing a steadily increasing drift current. Eventually the depletion region widens to where the internal electric field has risen to a value where the drift current magnitude equals the diffusion current magnitude. At this time a stable equilibrium condition is reached. From this time on, the depletion region width remains constant.

Fig. 1-10 plots the charge density, the electric field, and the potential versus distance across the junction. Starting with the charge density associated with the uncovered bound lattice ions in the depletion region on either side of the junction, we can find the electric field using the point form of Gauss' Law for the electric field, which states that the divergence of electric field E is equal to ρ/ε.

Because E varies only in the x direction, the divergence of E is dE/dx, and Gauss Law states that

$$\frac{d}{dx}E = \frac{\rho(x)}{\epsilon} \qquad (1 - 21)$$

Multiplying both sides by ***dx*** and integrating yields

$$E(x) = \frac{1}{\epsilon}\int \rho(x)\, dx \qquad (1 - 22)$$

Integrating the charge density distribution in Fig. 1-10 yields the plot of E versus distance x. Note that the E field is most intense at the junction. The potential V is the work done in moving a 1 C test charge through an E field, thus

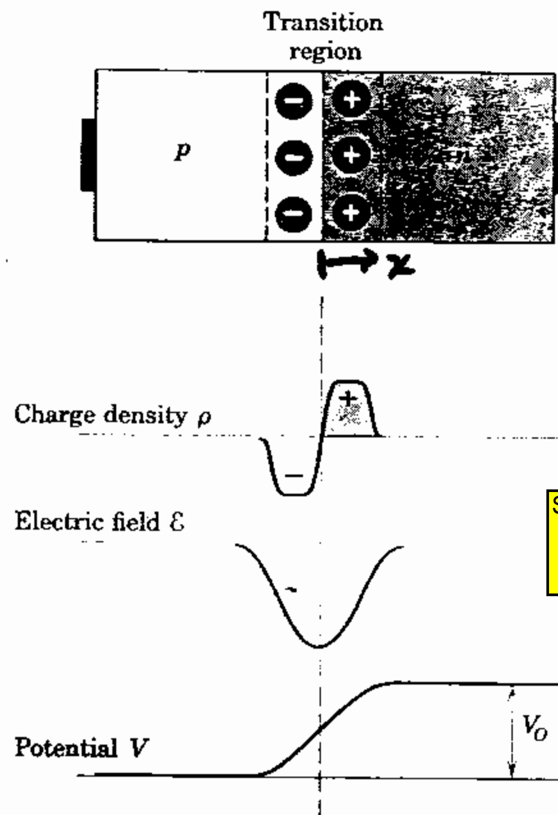
$$V(x) = -\int E(x)\, dx \qquad (1 - 23)$$

Integrating the E(x) distribution in Fig. 1-10 yields the plot of potential versus distance x.

The potential barrier, or "potential hill" of height Vo acts to oppose diffusion of majority carriers (holes to the right and electrons to the left) and to encourage the drift of minority carriers back across the junction.

Under open-circuit conditions, the net current is zero; the tendency of majority carriers to diffuse is just balanced by the tendency of minority carriers to drift across the junction as a result of the electric field. The height of the potential barrier, ***Vo***, is sometimes called the ***junction contact (or barrier) potential***. ***Vo*** can be calculated from the equilibrium condition, and is found to be a few tenths of a volt. However it cannot be used to cause external current flow, otherwise the principle of energy conservation would be violated; a pn junction ***cannot*** be used as a battery!

Fig. 1-10. Charge density, E field, and potential distribution across the open-circuited pn junction



$= V_{bi} = V_T \ln\left(\frac{N_a N_d}{n_i^2}\right)$

V_T (kT/q) = 26mV @ Room Temp

See Sedra and Smith Text

Eqn 3.2 and Eqn 3.48

Connecting an external conductor across the ***p*** and ***n*** regions creates two new metal-semiconductor contact potentials that just cancel the pn junction contact potential, ***Vo***.

A loosely analogous example might help us accept this claim. Consider a vertical column of gas. The molecules ***drift downward*** under the influence of gravity, but the they also ***diffuse upward*** because of their thermal energy. Therefore an equilibrium condition develops that is somewhat similar to that in an open-circuited pn junction. No net current gas flow occurs, because the average molecular drift rate downward is just canceled by the average molecular diffusion rate upward. A pressure difference (analogous to Vo) exists between the top and the bottom of the gas column, but this pressure difference ***CANNOT*** be used to cause any external gas flow in a second vertical column, since the same pressure difference would exist in the second gas column as well.

1.4.4 Forward-Biased pn Junction

The diffusion component of current is very sensitive to the barrier height, since only those majority carriers having kinetic energies in excess of qVo have sufficient energy to "climb" the potential hill and thereby diffuse across the junction.

The probability of a carrier possessing this amount of energy is

$$e^{\frac{-q\cdot Vo}{k\cdot T}} \qquad (1 - 24)$$

Placing an external voltage source across the diode with its positive terminal connected to the p region will cause the open-circuit equilibrium condition to be upset by decreasing the junction potential to a new value of (Vo - V), as shown in Fig. 1-11.