Experiment 13

THE SOLID STATE DIODE

REFERENCES:

General Appendix B: Fundamental Concepts of Thermal Physics

N. W. Aschroft and N. D. Mermin, *Solid State Physics* (1976, Thomson Learning, Inc.)

John L. Moll, *Physics of Semiconductors* (1964, McGraw-Hill, Inc.)

INTRODUCTION – THE PN JUNCTION DIODE

One of the simplest highly nonlinear electronic circuit components is the *semiconductor diode*. This *two-terminal* element behaves in a most asymmetric manner: its resistance is very low for currents flowing in one direction through the device, but it has an enormously high resistance to current flow in the opposite direction. Consequently, the diode acts as a "one-way valve" for electrical current (i.e., it is a *rectifier*). The most common type of diode is made from a semiconductor crystal divided into two layers with different impurity atoms mixed into it. The resulting structure creates a *PN junction* at the interface between the two layers which gives the diode its rectification property. The fascinating physics behind this behavior is very briefly introduced in this experiment's Appendix.

The schematic symbol of a diode is shown at right, along with a photo of a typical silicon signal diode. When the diode is forward-biased its resistance becomes very small, and current will flow through it in the direction shown (note that the schematic symbol includes an arrow (actually a triangle) which points in the direction of the current flow). Forward-biasing is accomplished by applying a voltage so that the diode's *anode* is at a more positive (+) voltage than its *cathode*, as shown in the figure. As you can see in the figure, the cathode is denoted by a line in the schematic symbol and is usually marked by a line or stripe on the physical diode's body. The anode consists of a metal terminal bonded to the *P-type* side of the semiconductor's PN junction; the cathode is bonded to its *N-type* side.

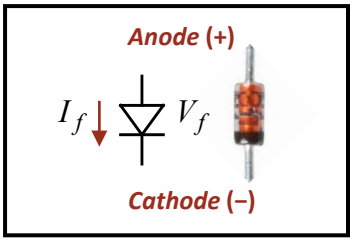

Figure 1: A typical silicon signal diode and its schematic symbol. The diode's glass case is actually only about 3mm long. When forward-biased, the diode can conduct large currents (as shown by the arrow).

The PN junction diode is very useful for constructing absolute value, peak detection, overvoltage protection, and more general nonlinear resistance circuits; as we will see, a diode's *current-voltage relationship* is actually exponential, so it is also useful as a component in exponential and logarithmic response amplifiers. The diode's characteristics are strongly temperature-dependent, so it also makes a sensitive, accurate temperature sensor. Even more importantly for much of today's technology, diodes can also emit and detect light (LEDs, laser diodes, and photodiodes), so the variety of modern applications of the seemingly simple diode is nearly endless.

THEORY – THE DIODE EQUATION

The semiconductor PN junction rectifier was invented by Russell Ohl at Bell Laboratories in 1939. Semiconductor device theory and fabrication was in its infancy at that time, and a detailed description of the physics behind his invention had to wait for William Shockley's 1949 theory (also at Bell Labs). Shockley's theory included the functioning of the *bipolar junction transistor*, which he invented in 1948. This device was a major improvement over the original working transistor invented at Bell Labs earlier that year by John Bardeen, Walter Brattain, and Shockley, for which they were awarded the 1956 Nobel Prize in physics.

A simple version of Shockley's theory in presented in this experiment's Appendix, which you should now review. The culmination of that discussion is his famous *diode equation*:

$$\boxed{I = I_R(e^{q_e V/\eta k_B T} - 1)} \tag{13.1}$$

where I and V are the PN junction current and voltage drop (positive when the junction is *forward-biased*, negative when *reverse-biased*), q_e is the magnitude of the electron charge, k_B is Boltzmann's constant, T is the absolute temperature, I_R is the diode's "ideal" reverse-bias *saturation current*, and η is the PN junction's *ideality coefficient* (which ranges in value from 1 to 2 depending on the diode's construction, among other factors). Note that except for the (–1) term (which ensures that the diode current vanishes if $V = 0$), the diode current I rises exponentially with increasing forward-bias voltage V.

The saturation current I_R depends on the energy gap of the semiconductor material used (1.12eV for silicon, 0.67eV for germanium at room temperature) as given by the following expression:

$$I_R = I_0\, e^{-E_g/\eta k_B T} \tag{13.2}$$

where I_0 is a constant. As explained in the Appendix, the diode's reverse-bias current, although very small, is actually significantly greater than I_R at large reverse-bias voltages; it is very temperature-dependent, as you would expect from the expression (13.2). Typical forward-biased and reverse-biased *IV characteristic curves* for a commercial silicon diode are illustrated in Figure 2.

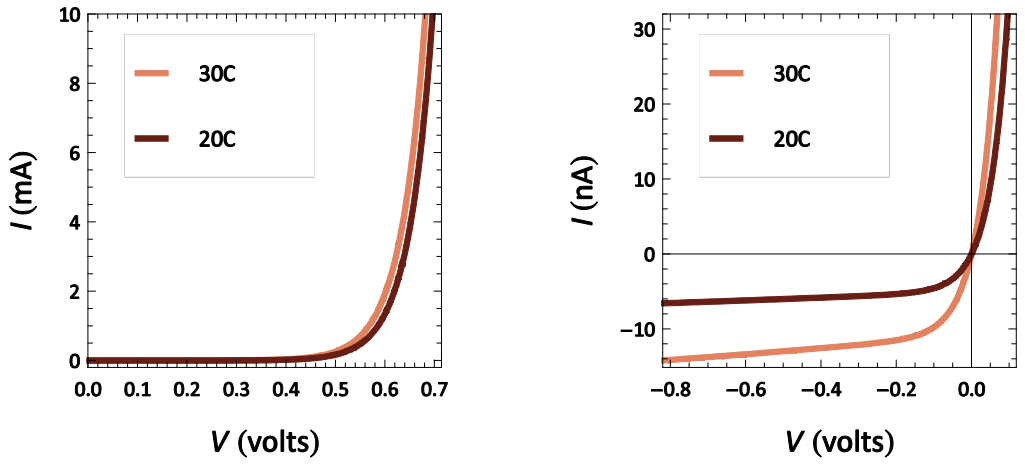

Figure 2: Forward-bias (left) and reverse-bias (right) IV characteristic curves for the 1N4148 silicon diode. Note the different vertical scales for the two plots (a factor of 10^6).

The purpose of this experiment is to investigate the physics of the semiconductor junction diode by testing the limits of the validity of the diode equations (13.1) and (13.2). You will measure the DC current-voltage (IV) characteristic relation for a forward-biased junction diode over a range of several orders of magnitude of current, at both room temperature and at the temperature of *dry ice* (194.67 K). You will examine a variety of PN junction diodes, including silicon and germanium types as well as *light-emitting diodes* (LEDs) of various colors and even a silicon *bipolar junction transistor* operating as an *ideal diode*.

EXPERIMENTAL SETUP

Since the current I through a PN junction is an exponential function of the junction voltage V, any noise in V could result in very large fluctuations in I. Thus it is wise to *set* the current I and then *measure* the diode voltage V which results. This is the method you will use.

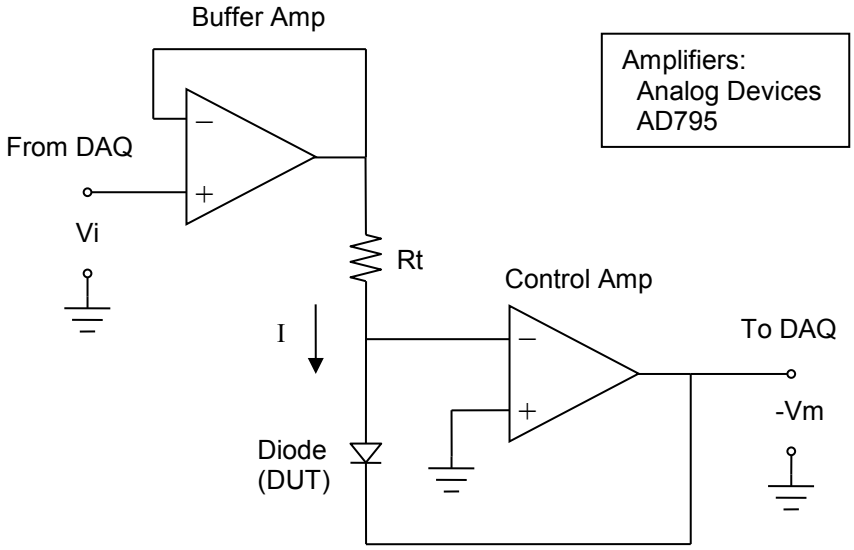

Figure 3: DAQ IV Control and Measurement Circuit

A special card is installed in the computer Data Acquisition (DAQ) interface box for this experiment. The computer will apply a control voltage to this card which will set the current I through the diode. The computer will then measure the voltage V across the diode for the established current condition. In order to keep the diode from heating up excessively because of the power dissipated in its small volume, the current will only be applied for a short time and then returned to zero.

A simplified diagram of the DAQ interface control circuitry used for this experiment is shown in Figure 3. The control and measurement process proceeds as follows:

1. The DAQ hardware's analog output channel supplies an input control voltage V_i to set the diode current I. This voltage is buffered by a precision amplifier which can supply up to 10mA of current at its output.

2. The current I flows from the buffer amplifier through the transfer resistor R_t and then through the diode. The current is absorbed by the output of the control amplifier. Two different transfer resistors are available, 1.0 kOhm and 2.0 MOhm (selected under computer control by a relay on the interface card).

3. The control amplifier continually and automatically adjusts its output voltage so that the voltage at the junction joining the diode and R_t is maintained at 0. As a result, its output voltage $-V_m$ is just the negative of the voltage across the diode V. This voltage is read by the DAQ hardware's analog input channel.

Because the control amplifier has a very large voltage gain (~10^5) and very small bias currents at its inputs (~1 pA) it maintains the voltage at the junction of R_t and the diode quite precisely at 0, and the current through R_t must also be the current I through the diode. Consequently $I = V_i / R_t$ and $V_m = V$. The DAQ control voltage has 16-bit resolution, so this setup can provide an accurately-set current over a range of less than 1 nA (10^{-9}A) to as high as 10 mA (10^{-2}A).

Small systematic errors remain, however, so an accurate calibration of the hardware should be performed in order acquire the best possible data when covering a 10^7 current range. The main sources of systematic error in the measurement hardware include:

- Offset voltage errors in the amplifiers and DAQ hardware (V_{io} and V_{mo})
- Actual value for the transfer resistor (R_t), rather than some nominal value
- Nonzero stray resistance in the wiring to the diode (R_s)

Including calibration constants which reflect these sources of error, we may write the relationships between the DAQ-commanded input voltage V_i, the diode current I, the diode voltage V, and the circuit measured output voltage V_m as follows:

$$\begin{aligned} I &= (V_i - V_{io}) / R_t \\ V_m - V_{mo} &= V + (I \times R_s) \end{aligned} \tag{13.3}$$

The calibration procedure will provide accurate estimates for the values of the calibration constants V_{io}, V_{mo}, R_t, and R_s. These values will then be used by the experiment's control software to calculate and save corrected values of the diode's I and V data using the equations (13.3).

EXPERIMENTAL PROCEDURE

The application program used to control the experiment is *DiodeIV*. The main window of the program is shown in Figure 4.

Also shown in the figure is a typical set of data collected to generate an IV characteristic curve for a silicon signal diode, the 1N4148 type. As you can see, the total current range of the displayed data extends from 200 pA to 10 mA (the high-current and low-current ranges overlap in approximately the 2−6 µA range). The nearly straight line on the semi-log plot joining the data points for currents above ~20 nA illustrates the remarkably accurate exponential relation between the voltage across and current through this semiconductor PN junction.

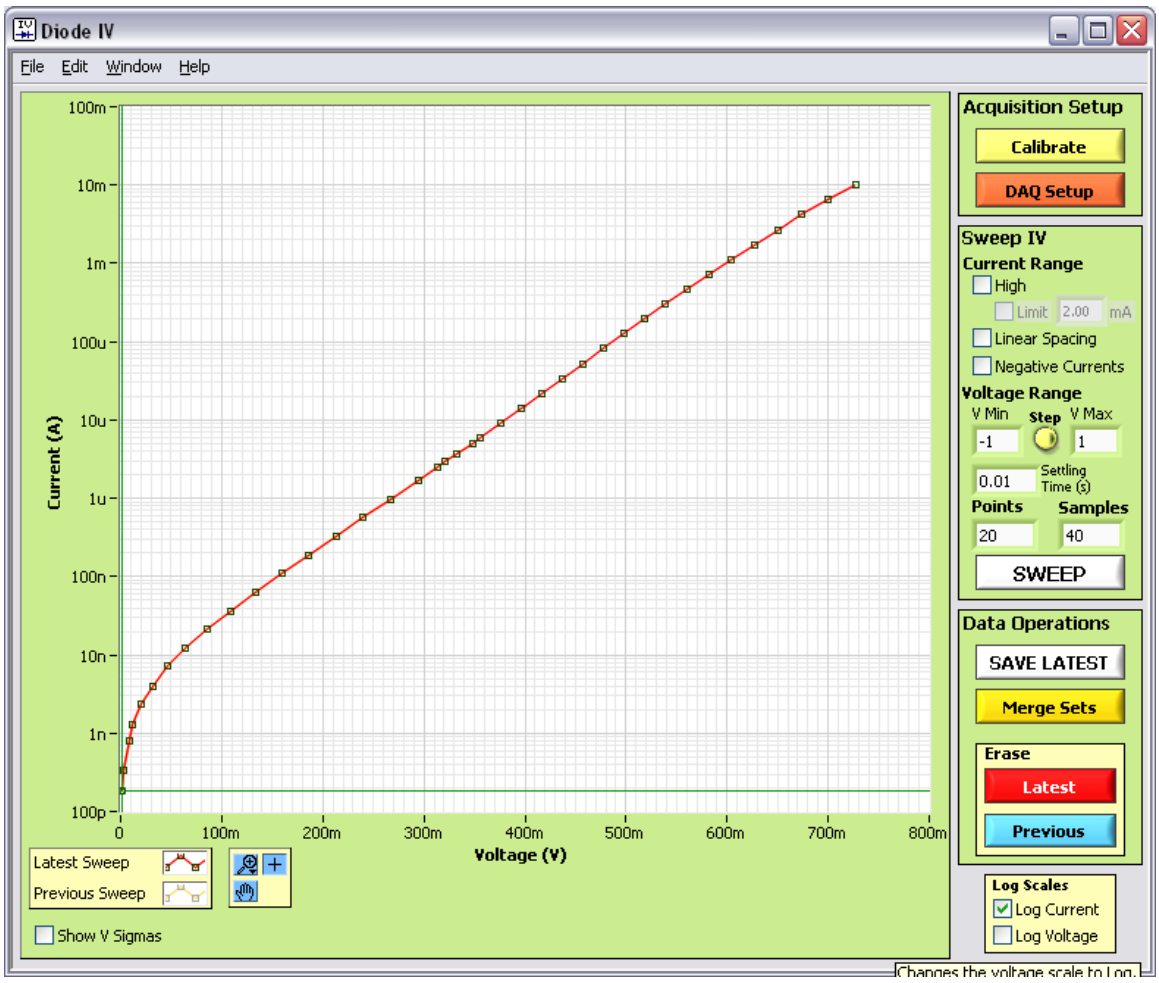

Figure 4: Diode IV Application Main Window. The displayed IV curve data is for a room-temperature silicon diode.

- ***Initial diode sweeps***

  Take initial room temperature IV sweeps of the silicon and germanium diodes and an *NPN transistor* (*collector* and *base* connected together and to the red test lead, *emitter* to the black test lead). The device and its connections to the diode IV circuit should be shielded from stray electric fields using a grounded metal can (why must the can be grounded to provide shielding?).

  Merge the data from both high-current and low-current sweeps of each device; since you haven't calibrated the hardware yet, the two sweeps should not quite join smoothly, but the data will be adequate for these initial observations. For each device (Si and Ge diodes and the transistor) use the displayed data to **estimate** the diode equation I_R and η values. The diode equation (13.1) is repeated below:

$$I = I_R(e^{q_e V/\eta k_B T} - 1) \tag{13.1}$$

  You should include images of these initial data sets along with your parameter estimates in your lab notebook.

- ***Calibration***

Following these preliminary diode data sweeps you are ready to calibrate the hardware.

Again examine equations (13.3) (repeated below) relating the input and measured circuit voltages, V_i and V_m, to the diode's current I and voltage drop V. The equations contain 4 calibration parameters which connect V_i and V_m to I and V.

$$\begin{aligned} I &= (V_i - V_{io})/R_t \\ V_m - V_{mo} &= V + (I \times R_s) \end{aligned} \qquad (13.3)$$

The purpose of the calibration of the apparatus is to determine accurate values for the four calibration parameters used in equations (13.3): V_{io}, V_{mo}, R_t, and R_s. These parameters might all depend on whether the high or low current range is selected, but in practice the only parameter which is most dependent on the selected current range is R_t (which, as mentioned previously, is the circuit element which is changed when you change the current range). The *DiodeIV* software includes a calibration window which allows you to set and adjust the values of the calibration coefficients; the window is shown in Figure 5.

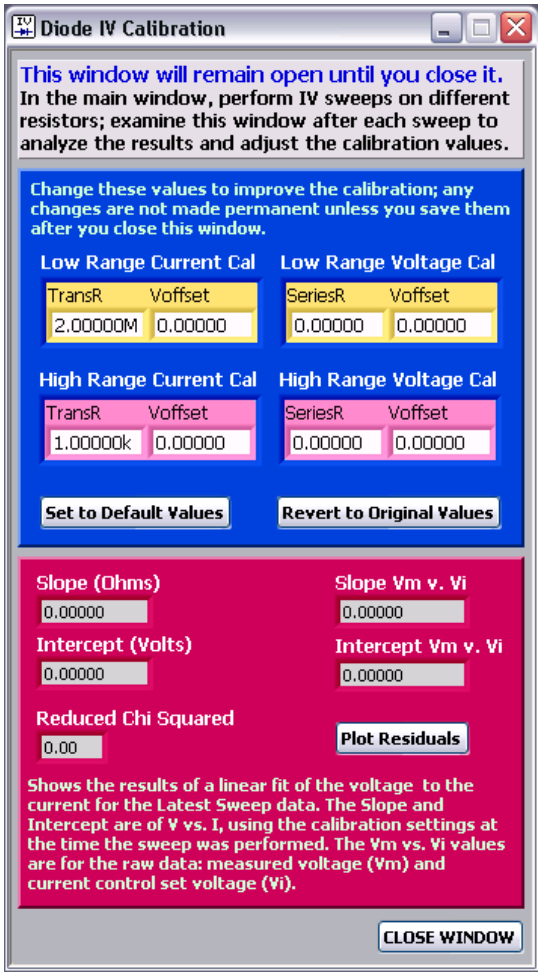

Figure 5: Diode IV Application Calibration Window. The upper (blue) area contains the currently set calibration parameter values; the lower (red) area displays calculated linear fit results to the latest IV sweep data.

You insert resistors of known value instead of a diode as the *device under test*. A resistor's linear IV relationship (Ohm's law, $V = IR$) implies that the resulting relationship between the circuit's input and output voltages, V_i and V_m, in equations (13.3) must also be linear. You take an IV sweep of a resistor, and the program then fits a straight line to the data, displaying the fit results in the lower (red) area of the calibration window. For each current range (high and low) there are 4 unknown calibration parameter values; taking IV sweeps of two different resistors will provide two

pairs of fit parameters (V_m v. V_i slope and intercept), so these 4 fit parameters (along with the measured values of the two resistors) may be used to solve equations (13.3) for estimates of the 4 calibration parameter values for a single current range. This is the "method" behind the experiment's calibration "madness."

Clever selection of the two calibration resistors can simplify the calibration process. Consider the second of equations (13.3): if the selected *R* is zero (or a tiny fraction of an Ohm), then $V = 0$ regardless of *I*. The two equations may then be combined to give:

$$(R = 0) \;\rightarrow\; V_m = \left(\frac{R_s}{R_t}\right) V_i + V_{mo} - \cancel{\frac{R_s}{R_t} V_{io}} \qquad (13.4)$$

Since the two amplifiers used in the circuit are the same type, we expect $|V_{io}| \sim |V_{mo}|$; additionally, we expect the stray resistance in the wiring to be small, so $R_s \ll R_t$. Thus we may safely discard the final term in (13.4), as shown. A linear fit to V_m v. V_i when $R = 0$ should thus have an intercept $= V_{mo}$ and a slope of R_s / R_t, giving the value of R_s (assume R_t has its nominal value, which is within 1% of its actual value).

The *DiodeIV* application should be configured for calibration data acquisitions as shown in Figure 6 on the next page. The two displayed IV sweeps in that figure were acquired with $R = 0$: the orange (Previous Sweep) data set was taken with default calibration settings (in particular, V_{io}, V_{mo}, and R_s all set to 0); the red (Latest Sweep) data set is also with $R = 0$, but V_{io} and R_s have been properly adjusted so that the result is a vertical line at $V = 0$ (except for random noise in the *V* measured data).

Now that two of the calibration parameters have been determined, next insert a resistor with $R \sim R_t$. The slope and intercept of V_m v. V_i in this case, along with equations (13.3) and the already determined values of V_{mo} and R_s, provide values for R_t and V_{io}:

$$(R \sim R_t) \;\rightarrow\; V_m = \left(\frac{R}{R_t}\right) V_i + \left(V_{mo} - \frac{R}{R_t} V_{io}\right) \qquad (13.5)$$

Evidently, $R_t = R / \langle \text{slope} \rangle$ and $V_{io} = (R_t / R) \times (V_{mo} - \langle \text{intercept} \rangle)$. These calibration steps should be accomplished for the high-current range using $R = 0$ and $R = 1\text{k}\Omega$. The values found for V_{io}, V_{mo}, and R_s are valid for both the high and low-current ranges; only R_t for the low-current range remains to be determined, which can then be done using $R = 180\text{k}\Omega$.

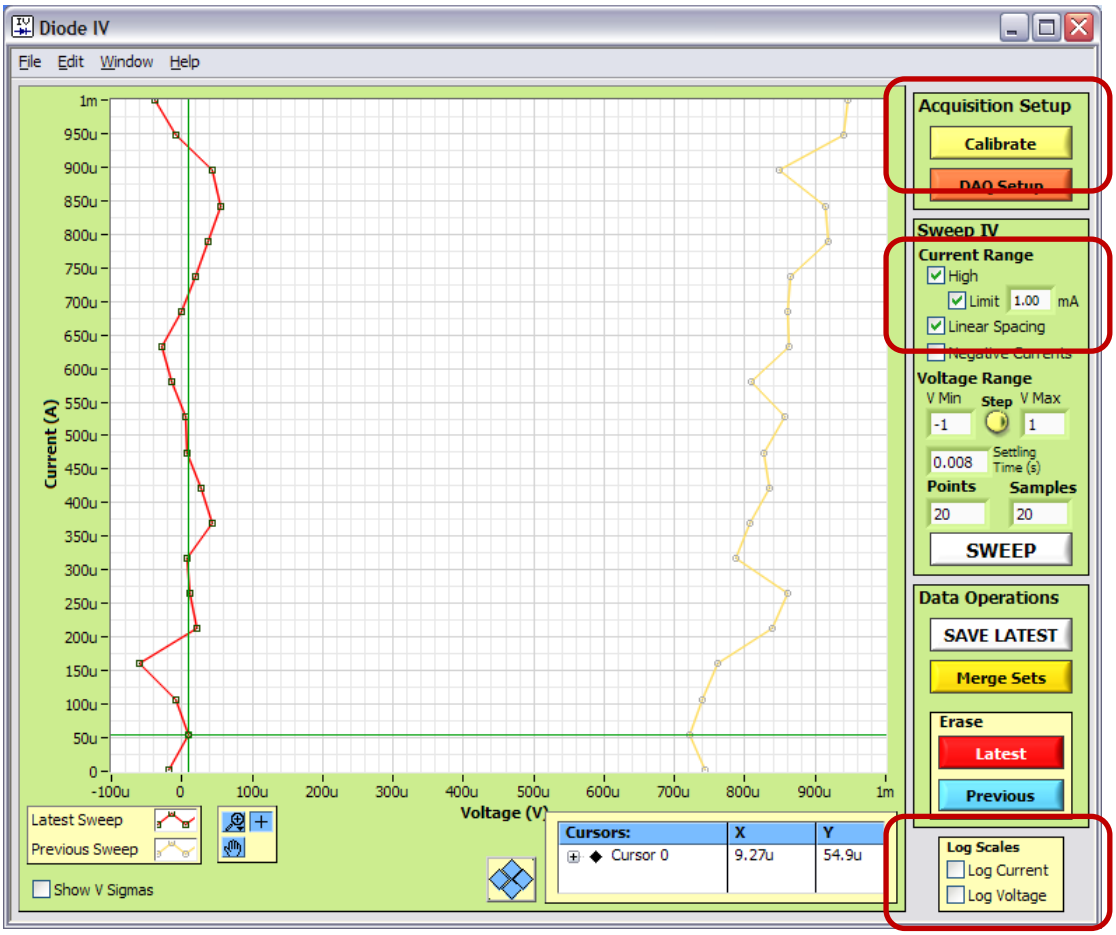

Figure 6: Diode IV configured for calibration IV sweeps (the red boxes highlight the controls which must be set).

(1) The *Calibrate* button opens the calibration window (Figure 5).

(2) The *High* current range must have its maximum current limited so that the measured voltage won't exceed the selected voltage range *VMax* (e.g., 1mA for a 1kΩ calibration resistor, as shown in the figure).

(3) *Linear Spacing* of the data point currents is appropriate for a resistor IV sweep.

(4) *Log Scales* are inappropriate for a linear IV sweep of a resistor, so both boxes should be unchecked.

To summarize the calibration steps, you must:

1. Set up the program for calibration (Figure 6) and set the calibration parameters to their default values (button in Figure 5).
2. Short the sample wires so that $R = 0$ and take IV ***high current*** range IV sweeps. Determine values for V_{mo} and R_s; enter the results into the appropriate boxes for both the ***Low Range Voltage Cal*** and ***High Range Voltage Cal*** (Figure 5).
3. Accurately measure the resistance of a test resistor with a value of approximately $R = 1\text{k}\Omega$ and connect it to the circuit. Make sure the ***high range current limit*** is set appropriately (Figure 6). Use ***high current IV*** sweep data to determine an accurate value for the ***High Range Current Cal*** R_t (Figure 5). Also determine a value for V_{io} and enter it into both the ***High Range Current Cal*** and ***Low Range Current Cal*** offset boxes.
4. Accurately measure the resistance of a test resistor with a value of approximately $R = 180\,\text{k}\Omega$ and connect it to the circuit. Use ***low current IV*** sweep data to determine an accurate value for the ***Low Range Current Cal*** R_t (Figure 5). This should complete the calibration process.

- ***Calibrated data collection***

After completing the calibration, the data acquisition and display settings should be returned to values appropriate for diode IV data acquisition as shown in Figure 4 (no high current limiting, linear spacing deselected, log current scale).

Repeat the room temperature IV sweeps of the silicon and germanium diodes and an *NPN transistor* (*collector* and *base* connected together and to the red test lead, *emitter* to the black test lead). If the calibration was properly completed, the high-current and low-current data sets should join smoothly with only a very small offset where the sets overlap (as is the case in the Figure 4 data plot). If this is not the case, then the most likely culprit is the calculated value for V_{io}. Double-check the sign of this value and adjust it by a few 10's of microvolts if necessary to improve the overlap offset of your diode IV data.

Now cool each device by burying it in dry ice and repeat the IV sweeps. The cup containing the dry ice should be placed in the shielding can to avoid interference from external fields. Do not bury the wire connector clips in the dry ice, however. The device will have stabilized at the dry ice temperature (194.67 K) when multiple IV sweeps have settled down into a repeatable data set.

Estimate a new I_R value for the cold germanium diode and compare it to your original, room temperature estimate. Use these two values to estimate E_g, the germanium gap energy (in eV). Recall the formula for I_R, equation (13.2), repeated below:

$$I_R = I_0 e^{-E_g/\eta k_B T} \tag{13.2}$$

All calibrated sweep data (room temperature and cold) for each diode type (including the transistor) should be saved for later analysis.

- ***Additional observations***

If you have time, obtain IV sweep data for the variously colored LEDs available in the lab. You will need to increase the *DiodeIV* application's voltage range by actuating the round yellow button between the *V Min* and *V Max* windows (Figure 4).

ANALYSIS

Use your data to evaluate the accuracy of the simple PN junction theory, equations (13.1) and (13.2). *CurveFit* includes a diode equation fitting function; you should initially attempt to fit the entire merged data set (high and low current ranges) of each device and provide a thoughtful analysis of the fit residuals. You may then find it wise to attempt diode equation or exponential fits to various subsets of the data to investigate the variation of the ideality parameter η for each device. Which device's behavior appears to match the theory most closely (silicon, germanium, or the transistor)?

It is possible that your data may have evidence of additional series resistance effects when the applied currents are large (in the milliamp range), since the diode itself will have some finite resistance between each of the device's connecting leads and its internal PN junction. How would this additional resistance show up in the measured data? Assuming that this resistance changes only slowly with temperature, then would you expect it to have a greater effect on the cold data or the room temperature data?

Using your best estimates of the room temperature and cold values of I_R and η for each device, determine E_g using equation (13.2). How do your results compare to the (room temperature) semiconductor gap energies of $1.12\,\text{eV}$ for silicon and $0.67\,\text{eV}$ for germanium (the transistor is also silicon)? Actually, a semiconductor's band gap energy is slightly temperature-dependent, increasing very nearly linearly with decreasing temperature over the temperature range you are investigating. Both semiconductors' E_g values increase by about $0.03\,\text{eV}$ as temperature decreases by $100\,\text{K}$.

The PN junction theory presented in the Appendix assumes a very simple, one-dimensional model for a PN junction diode device. In what ways do you think that this model might be an oversimplification of a real device?

PRELAB PROBLEMS

1. Examine the room-temperature silicon diode data set plotted in Figure 4 (page 13-6) and answer the following questions using that data.
   a. Extrapolate the exponential region of the data (straight line in the plot) to 0V. What is an estimate of the current at this 0V intercept? How does this current value relate to the diode's I_R (equation (13.1))?
   b. Estimate the voltages where the curve passes through $I = 100\,\text{nA}$ and $I = 1\,\text{mA}$. Use the voltage difference and the current ratio (10^4) to estimate the value of η, also from equation (13.1).
   c. What feature of the plot shows the strongest evidence of the (−1) term in (13.1)?

2. Reexamine the silicon diode room-temperature data in Figure 4 (page 13-6) and your answers to question 1. What would you estimate the value of I_R to become when the diode is cooled to dry ice temperature (194.67 K)? Sketch the two IV curves (room temperature and dry ice) on a single log-current, linear-voltage plot. Ignoring the (−1) term in the diode equation, at what voltage should the two curves intersect (assume that silicon $E_g = 1.12\,\text{eV}$)?

3. You are calibrating the high-current range using a test resistor with a value of $R = 997.1\,\Omega$, and the nominal value of $R_t = 1.000\,\text{k}\Omega$ is set in the *DiodeIV* calibration window (Figure 5 on page 13-8). An IV data sweep results in a V_m v. V_i slope of 0.9993. What should be the corrected value for R_t?

4. Estimate the width of a silicon PN junction depletion layer (this may be a tough one). Refer to Figure A-3 on page 13-21 and assume that the P-type acceptor and N-type donor concentrations are uniform on either side of the junction at 5×10^{16} per cm^3, so that the volume charge density ρ of the ions on either side of the depletion layer has a magnitude of 8×10^3 Coulomb per m^3. Assume that the edges of the depletion layer are abrupt and that outside the layer the electric field vanishes, as shown in the figure. Clearly, the total net ion charge within the depletion layer must be 0, and the layer extends for equal distances on either side of the PN interface (since the ion concentrations on the two sides of the interface are equal). The permittivity of silicon is $\varepsilon = 11.7\,\varepsilon_0$.
   a. Let z measure the distance into the depletion layer from its left-hand boundary in the P-type side (Figure A-3). Show that the electric field strength increases linearly as $|dE/dz| = |\rho|/\varepsilon \approx 7.7\times10^3\,(\text{V/m})/\text{Å}$ approaching the PN interface.
   b. If $V_j = 0.9\,\text{V}$ is the equilibrium potential drop across the depletion layer, show that the total width of the depletion layer is approximately $2200\,\text{Å}$, and that the electric field strength at the PN interface is approximately $8\times10^6\,\text{V/m}$.

APPENDIX – ELEMENTARY PHYSICS OF THE PN JUNCTION DIODE

- ***Insulators, conductors, and semiconductors***

The electrical conductivities of solid materials for the most part fall into one of two classes: conductors and insulators (metals make up most of the conductors, and nonmetals are usually insulators). Although all materials are very nearly electrically neutral (equal numbers of protons and electrons, so that they carry no net charge), the nature of the chemical bonds which bind the atoms or molecules of a solid to one another determines its class of electrical conductivity.

An atom's *valence electrons* — the outer, most weakly bound electrons — are the ones which participate in chemical bonding. The atomic nucleus along with the much more strongly bound inner electrons comprise a positively-charged *ion core* which remains intact and is surrounded by the interacting valence electrons. In a solid the chemical bonding process causes these many ion cores to arrange themselves in a mostly regular, crystalline structure. This regular, periodic array of positively-charged cores creates a similarly regular, periodic electrostatic field within which the myriad valence electrons move.

The quantum-mechanical nature of these microscopic, negatively-charged particles (the valence electrons) as they experience the periodic electrostatic potential of the ion cores requires that they each occupy a state of motion (and total energy) in one of several distinct *energy bands*, analogous to the quantized energy states an electron may occupy in a single atom or molecule. The width of a typical energy band is on the order of a few to several electron volts (same order of magnitude as the binding energy of a valence electron in one of the atoms), and adjacent energy bands are often separated by a similar energy, although they may also overlap (depending in a complicated way on the geometry of the crystal structure and the nature of the material's interatomic bonds). Each band has enough distinct quantum states to contain twice the number of electrons as there are molecules in the macroscopic solid crystal (i.e. $\sim 10^{22}$–$10^{23}/\text{cm}^3$).

Room temperature ($\approx 290\,\text{K}$) corresponds to random, thermal particle energies of $\sim k_B T \approx (1/40)\,\text{eV}$ (electron volt), much smaller than the width of an energy band but much larger than the energy spacing between the individual states in a band. Because electrons are subject to *Pauli Exclusion* (each electron must be in a unique, distinct quantum state), the valence electrons of all the various atoms in a solid fill the available states starting with the lowest available energy. Because room temperature corresponds to a fairly small energy, the energy of the topmost occupied states is fairly well-defined and is called the electrons' *Fermi energy*. Random thermal jostling can only affect the states of individual electrons with energies near the Fermi energy, because those with much

lower energies are surrounded by quantum states already occupied by other electrons, so they're stuck in their current states.

Now, one of two situations can occur for our valence electrons in a solid:

1. The number of electrons is such that they exactly fill all the available states in some number of energy bands, and higher energy bands are completely empty.

2. One energy band (or possibly more, if some bands overlap) is only partially filled and has many unoccupied states still available; all other bands are either completely filled or completely empty.

Electrons occupying a completely filled energy band do not participate in electrical conduction. The reason for this is that such a band corresponds to all physically possible states of individual electron motion in all directions consistent with the energies of the electrons in that band. Applying an external electric field doesn't change this situation unless the field is so intense that it can cause electrons to transition to another (partially filled or empty) energy band. Thus, no new net motion of electrons can be induced by the presence of the field, so the electrical conductivity contributed by a completely full (or, of course, completely empty) energy band is zero.

This last result implies that solids with situation (1) above are *insulators* (or maybe semiconductors). Since each band has twice the number of states as there are molecules in the crystal, insulating materials most often arise when there is an even number of valence electrons participating in the chemical bonding forming the solid. Situation (2), on the other hand, allows electrical conduction to proceed using the electrons in the partially-filled energy band. Electrons near the Fermi energy in the band have a wide selection of nearby empty states, so an applied electric field can accelerate them, and their resulting motions can carry a net flow of charge (electric current) through the solid. These materials are *conductors*, and partially-filled energy bands are characteristic of the so-called *metallic bond*.

Semiconductors have valence electrons whose situation falls into category (1): bands containing electrons are completely filled, at least at cold temperatures. What makes them different from insulators, however, is that the bottom of the nearest empty energy band (called the *conduction band*) is only about an eV or so away from the top of the highest-energy filled band (the *valence band*). Consequently, random thermal jostling of the ions in the lattice can occasionally impart enough energy to an electron with an energy near the top of the valence band to excite it into a level near the bottom of the conduction band. In this case both the valence band and the conduction band become *partially* occupied (although just barely), and the material becomes a poor conductor (poor because only a tiny fraction of the valence electrons get bumped up into the conduction

band). The higher the temperature, the greater the number of valence electrons thermally excited into the conduction band — the number goes as:

$$n_i \propto T^{3/2} e^{-E_g/(2k_BT)} \tag{13.A.1}$$

where E_g is the magnitude of the energy gap between the valence and conduction bands and k_B is Boltzmann's constant. In the case of silicon, this amounts to $\sim 10^9$ electrons per cm^3 at room temperature (compare with copper's $\sim 10^{23}$ per cm^3). The derivation of this expression is an interesting exercise in statistical mechanics and is provided in the last section of this Appendix.

The archetypal semiconductors are the elements silicon ($E_g = 1.12\,\text{eV}$), and germanium ($E_g = 0.67\,\text{eV}$), each of which form a diamond crystal lattice with four *covalent bonds* per atom. Carbon in its diamond form ($E_g = 5.5\,\text{eV}$) is beginning to find applications in solid-state devices, but its large energy gap makes it more properly classified as an insulator. Several compounds and alloys form commercially important semiconductors, including GaAs, InP, GaAsP, and InGaN.

- ***Electrons and holes; impurities and doping***

The diagram at right illustrates the distribution of electrons between the top of the valence band and the bottom of the conduction band for a semiconductor at a fairly high temperature (so that there have been a considerable number of electrons excited into the conduction band). The density of the quantum states in each band grows as $\sqrt{\Delta E}$ as you move away from the band edges, as shown by the right-hand curves in the figure. The Boltzmann factor $\exp(-\Delta E/k_BT)$ gives the relative probability that any one state is occupied in the conduction band or unoccupied in the valence band (see General Appendix B, ***Fundamental Concepts of Thermal Physics*** for a brief review of the statistical mechanics principles we use here). Because the number densities of the conduction electrons and the holes in their respective bands are low (much smaller than the crystal's atomic number density) and the spacing between states in a band is much smaller than k_BT, the charge

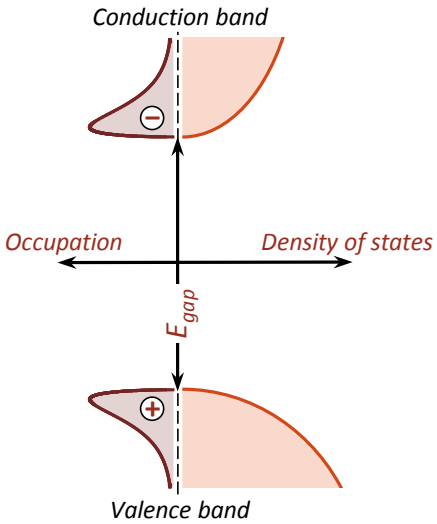

Figure A-1: Densities of states and occupations by electrons (conduction band) and holes (valence band) for a pure semiconductor (*intrinsic* charge carriers only); energy increases in the vertical direction. Because the density of the charge carriers is small the kinetic energy distribution of the charge carriers in each band is classical.

carriers will distribute themselves in accordance with the classical Maxwell distribution, so the overall occupation densities go as the left-hand curves in Figure A-1.

The dynamics of the relatively small number of electrons in the conduction band is quite accurately approximated by treating them as classical particles (with a negative charge of $-q_e$, of course), but their *effective mass* is determined by the shape of the density of states curve near the bottom of the conduction band: the sharper the curve near the minimum, the lighter the effective mass. In the valence band only a small fraction of the states near its top are unoccupied. Interestingly, the dynamics of the remaining electrons near the top of the valence band are such that they have a *negative effective mass*, since the density of states *decreases* with increasing energy near the top of the band.

The consequence of this unusual electron behavior near the top of the valence band is that the unoccupied quantum states evolve as though they were *positively charged particles* $(+q_e)$ *with a positive effective mass and with energies increasing as they move further down from the band top*! These "positive charge carriers" near the top of the valence band are called *holes*. In fact, the holes in a semiconductor valence band are *completely equivalent* to "real" particles such as the conduction band electrons, so *their particle nature is just as valid*. Thus, when an electron is excited from the valence band to the conduction band, *two* charge carriers are created: the electron $(-q_e)$ and the hole it left behind $(+q_e)$. Since the energy an electron must gain to cross the energy gap between the valence and conduction bands is at least E_g, but two particles were created by this transition, *the required energy per particle* is $E_g/2$ — this is a convenient "hand-waving" explanation of the extra factor of 2 in the Boltzmann expression (13.A.1).

The conductivity of a semiconductor is proportional to the volume density of its charge carriers (electrons + holes); thus a pure semiconductor has a conductivity which is a very strong function of temperature, rising rapidly as temperature increases, as indicated by expression (13.A.1). This effect is used to make a *thermistor*: a resistor with a large, *negative temperature coefficient* (decreasing resistance as temperature rises) which acts as a very sensitive, fast-acting temperature sensor for the range of about −100°C to +150°C.

Semiconductor materials are custom-made to be much more flexible and useful through the process of *doping*: introducing various amounts of impurity atoms into the semiconductor crystal which have a different valence than the semiconductor. For example, mixing a small amount of phosphorous (valence 5) into a silicon crystal will introduce a random distribution of atoms each with an extra valence electron left over after it forms bonds with surrounding silicon atoms. What would be the consequences of these extra electrons to the physics of the material? It turns out that the energy of this extra valence electron is very close to the energy of the bottom of the conduction band (in

the case of P in Si, the energy is only 0.044 eV below the conduction band). If there are relatively few of these *donor* impurity atoms, then it is very likely that such electrons will eventually be excited into the conduction band by the thermal motions of the ions: once there they quickly drift away from their parent impurity atoms and are unlikely to recombine with them. So even if the ambient temperature is cool enough that almost no electrons would be excited from the valence band to the conduction band, electrons from donor impurities will nearly all eventually find their way into the conduction band, providing a largely temperature-independent cadre of negative charge carriers $(-q_e)$ along with the same number of fixed, positively-charged ions distributed throughout the crystal lattice. Such a material is called an *N-type semiconductor*.

Similarly, introducing a valence 3 impurity atom (such as aluminum into silicon) will leave an unsatisfied bond because of the missing electron. Again, the energy required to promote a nearby silicon valence electron into this spot is small compared to the semiconductor's energy gap (0.057 eV for Al in Si). Thermal agitation will eventually do the trick, and the vacated valence state becomes a hole which quickly drifts away from the impurity atom, trapping the promoted electron at the impurity site. Thus these *acceptor* impurity atoms become fixed, negatively-charged ions in the lattice, whereas an equal number of holes form a nearly temperature-independent group of positive charge carriers $(+q_e)$, creating a *P-type semiconductor*.

Adding dopants to a semiconductor can not only introduce charge carriers (called *extrinsic charge carriers*), but will also suppress the thermal creation of electron-hole pairs described by equation (13.A.1), called *intrinsic charge carriers*. This is because the product of the number of conduction electrons (n_c) and the number of holes (p_v) is related to the number of intrinsic charge carriers that would be thermally created in a pure (undoped) semiconductor (n_i) by the laws of statistical mechanics:

$$n_c\, p_v \;=\; n_i^{\,2} \tag{13.A.2}$$

(this expression is also derived in the final section of this Appendix). For example, the addition of 1 part per million phosphorous to a silicon crystal would introduce 5×10^{16} extrinsic conduction electrons per cm^3; with $n_i \sim 10^9$ electrons per cm^3, we see that there will be only $p_v \sim 100$ holes per cm^3! These holes are called *minority carriers* in the N-type silicon under discussion; the conduction electrons are the *majority carriers*. Since for this example $n_i \ll n_c$, the temperature dependence of n_c will be quite small, so, from equation (13.A.2), $p_v \propto n_i^{\,2}$. Thus *the temperature dependence of the minority carriers is very large*: from equation (13.A.1),

$$\text{<minority carrier density>} \;\propto\; T^3\, e^{-E_g/(k_B T)} \tag{13.A.3}$$

- ***The equilibrium PN junction***

Now consider the case of a semiconductor crystal with *inhomogeneous doping*. As a concrete (but quite artificial) example, assume that we take a single P-type crystal and a single N-type crystal and join them along a planar boundary so as to form a single crystal with an abrupt change in doping at this boundary. The result is a *PN junction* (Figure A-2) at the interface between the two semiconductor types.

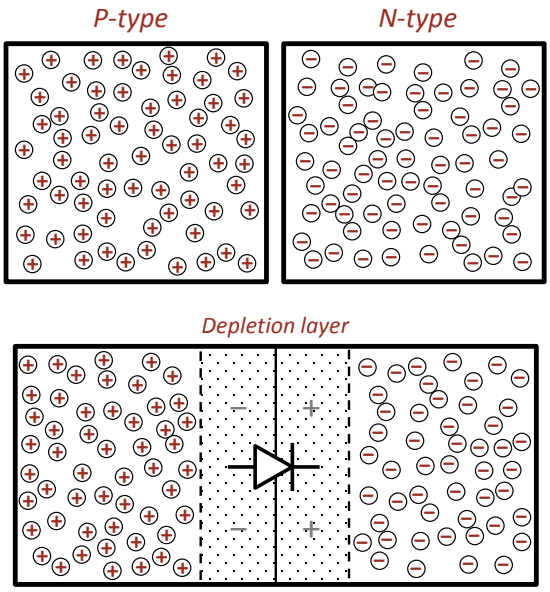

Figure A-2: Formation of a PN junction diode and its depletion layer.

Far from the boundary the charge carrier densities must approach their homogeneous, thermal equilibrium values. Near the interface, on the other hand, the large gradients in the hole and conduction electron densities will drive diffusion of these charge carriers across the boundary, where they will eventually recombine with carriers of the opposite sign. The reduced majority carrier densities near the boundary induce a net charge density and resulting electric field near the interface because of the now unbalanced charges of the impurity ions in each semiconductor. This electric field will repel the majority charge carriers on either side of the boundary, and an equilibrium condition is reached preventing further net diffusion of carriers across the boundary (bottom illustration in Figure A-2). It turns out that the equilibrium situation will be attained only when a region near the PN interface is almost completely devoid of charge carriers: the so-called *depletion layer*, as shown in Figure A-2. The charge density in this region is then given by the number densities of the impurity ions on each side of the boundary, which are nearly equal to the corresponding majority carrier number densities far from the boundary.

The electric field near the boundary generates a potential difference between the P-type and N-type sides of the junction, with the N-type material at the higher potential. If the impurity concentrations are each very much greater than the semiconductor's intrinsic carrier density (n_i), then the equilibrium *contact potential* (V_j) of the PN junction will be a significant fraction of the semiconductor's *gap voltage*: $E_g / q_e \equiv V_g$ (1.12 V for silicon). The result, as we shall see, is the creation of the *PN junction diode*.

Consider, for example, a silicon PN junction diode in which the N-type's donor concentration and the P-type's acceptor concentration are approximately equal. Once the

equilibrium depletion layer is established, the electric fields far from the layer will vanish, so the total number of unbalanced donor and acceptor ion charges on either side of the layer must be equal and opposite. For simplicity's sake, assume that the impurity concentration on either side of the PN interface is uniform, and that the depletion layer boundaries within the P and N type sides are abrupt and well-defined. In this case the electric field and the potential will vary across the depletion layer as illustrated in Figure A-3.

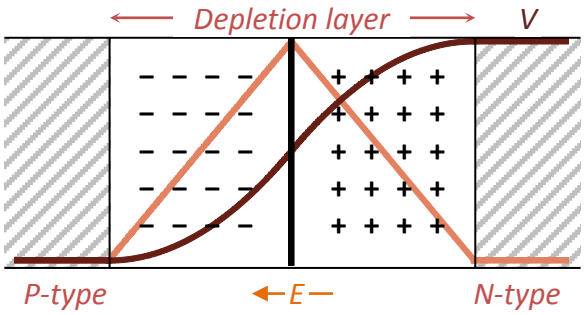

Figure A-3: Approximate variation in the electric field strength *E* (lighter curve) and resulting potential *V* (darker curve) across a PN junction's depletion layer. The electric field within the depletion layer points from the N-type material toward the P-type side (right-to-left, as shown by the arrow); if the impurity concentrations are large, then the equilibrium contact potential difference across the depletion layer will be a large fraction of the semiconductor's gap voltage V_g, and the device will be an effective diode.

Since the potential changes by $V_j \lesssim V_g$ across the depletion layer, it is straightforward to estimate its equilibrium width, which will typically be in the range of $10^2 - 10^4$ Å (about 2200 Å for silicon with a part per million doping), and the magnitude of the electric field at the interface is in the range of $10^5 - 10^7$ V/m.

- ***The PN junction I-V characteristic curve***

The equilibrium configuration (Figures A-2 and A-3) is maintained when the rate that the holes diffuse into the depletion layer from the P-type side (against the contact potential V_j) matches the small rate of hole diffusion from the N-type side (where holes are the minority carriers), so that the net flow of holes across the junction is 0; a similar condition holds for the conduction electron diffusion at the junction.

Holes entering the depletion layer from the N-type side are not impeded by the presence of the contact potential — on the contrary, the electric field in the depletion layer will accelerate them through it to the P-type side. This implies that the rate of the minority hole diffusion will simply be proportional to the hole density on the N-type side, which is given in relation (13.A.3) to be proportional to $e^{-E_g/(k_B T)}$, and similarly for the

minority electron diffusion from the P-type side. The majority carriers must cross the barrier imposed by the junction potential (V_j), so only those carriers with kinetic energies larger than $q_e V_j$ can cross to the other side; the number of such energetic carriers will be proportional to the Boltzmann factor $e^{-q_e V_j/(k_B T)}$ (because their kinetic energy distributions are classical, as mentioned before). At equilibrium, these two rates match; this requirement determines the value of the equilibrium junction potential V_j.

When an external bias voltage V is applied across the PN junction, this applied potential will reduce the junction contact potential by the same amount: $V_j' = V_j - V$ ($V > 0$ is forward-biased). As a consequence, the fraction of majority carriers with energies high enough to surmount the contact potential is now proportional to $e^{-q_e(V_j - V)/(k_B T)}$, a factor of $e^{q_e V/(k_B T)}$ times as large. Thus there will be a net current flow across the junction given by the difference between this new rate and the equilibrium rate, which was originally proportional to $e^{-E_g/(k_B T)}$:

$$I \propto e^{q_e V/k_B T} e^{-E_g/k_B T} - e^{-E_g/k_B T} = e^{-E_g/k_B T}(e^{q_e V/k_B T} - 1)$$

This simple result is known as the *ideal diode equation*:

$$I = I_R(e^{q_e V/k_B T} - 1); \quad I_R \equiv I_0 e^{-E_g/k_B T} \tag{13.A.4}$$

V is the applied bias voltage (+ for forward-bias), V_g is the semiconductor's gap voltage, I_0 is some constant, and I_R is the diode's asymptotic *reverse leakage current*. Thus, the ideal diode's forward current rises exponentially with forward bias voltage (for voltages of more than a few tens of millivolts), and has some small, temperature-dependent leakage current when reverse-biased ($V < 0$).

The above equation is not quite right, because its derivation ignores an effect which is especially important for the behavior of a silicon diode: generation and recombination of charge carrier pairs in the depletion layer. The assumption in the argument leading up to the diode equation was that the only charge carriers present in the depletion layer entered it through diffusion from the regions outside the layer, and that all of these carriers pass through the depletion layer without interacting. Actually, thermal excitation of electron-hole pairs will occur in the depletion layer, just as it would in a pure semiconductor; similarly, recombination of electrons and holes may also occur within the depletion layer among those charge carriers which diffuse into it. Generally, because of this depletion layer recombination process, the number of charge carriers entering the depletion layer of a forward-biased PN junction may be significantly larger than the number which escape, especially for small forward-bias currents in relatively large E_g diodes such a silicon.

The depletion layer generation and recombination processes also depend exponentially on temperature, but the exponent goes as $q_e/2k_BT$ rather than q_e/k_BT as in (13.A.4). The combination of this process with the ideal diode process leads to a "slight" modification of the ideal diode equation:

Diode equation

$$I = I_R(e^{q_e V/\eta k_B T} - 1); \quad I_R = I_0 e^{-E_g/\eta k_B T} \tag{13.A.5}$$

The *ideality coefficient* η depends on the importance of the depletion layer recombination process; it is a weak function of I and T and ranges between 1 and 2. For example, the 1N4148 silicon diode has $\eta \approx 1.9$ and $I_R \approx 5\,\text{nA}$; thus $q_e/\eta k_B T \approx 20\,\text{volt}^{-1}$ at 20°C. The exponential dependence of I on forward-bias voltage V is obviously strongly temperature-dependent. Figure A-4 (left) schematically illustrates the charge carrier behavior in a forward-biased PN junction. At large forward-bias voltages (and currents), the depletion layer is much narrower than its equilibrium width, and charge carrier recombination within this narrower layer is less likely. Consequently, the ideality coefficient η may

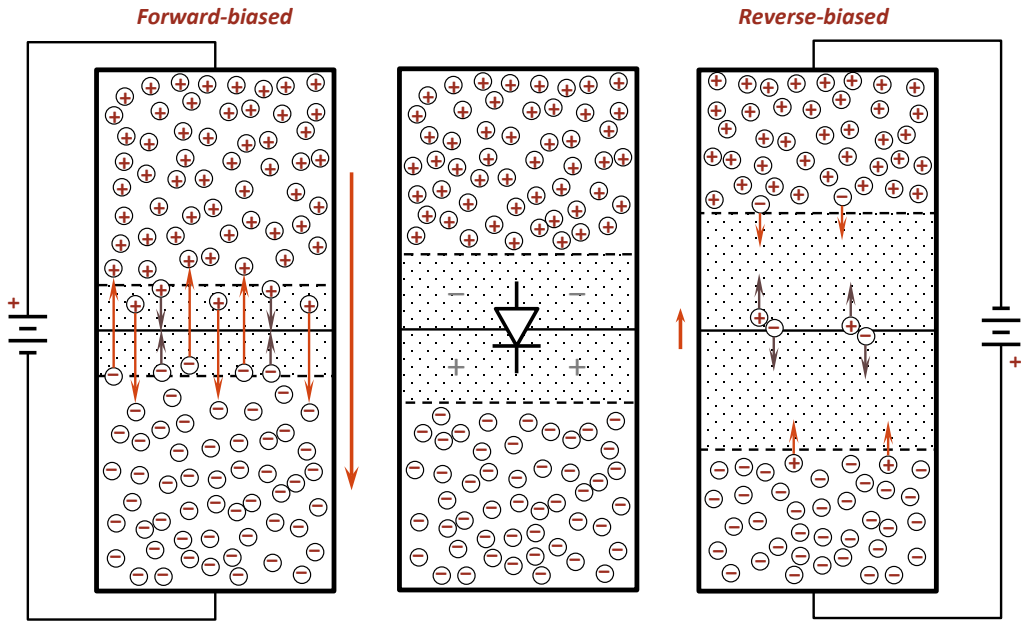

Figure A-4: Depletion layer width and charge carrier diffusion of a PN junction as affected by applied bias voltage. Forward-bias (left) reduces the height of the potential barrier to majority carrier diffusion and decreases the depletion layer width, so many majority carriers can diffuse into and through the depletion layer; minority carrier diffusion is largely unaffected. Carriers that cross the depletion layer and recombine with majority carriers on the opposite side are indicated by the orange arrows; those that recombine inside the depletion layer are shown with gray arrows. Reverse-bias current is completely dominated by *minority carrier diffusion*: those that enter the depletion layer from the bulk semiconductor (orange arrows) and those pairs that are thermally generated within the depletion layer (gray arrows).

approach 1 at high forward-bias currents.

When the junction is reverse-biased (Figure A-4, right-hand figure), the depletion layer size grows roughly as $\sqrt{1+V/V_j}$ (V is the reverse-bias voltage). The total thermal electron-hole generation rate within the depletion layer is proportional to its size, so the reverse current due to this generation does not "saturate" at the I_R value given by equation (13.A.5), but continues to grow slowly with increasing reverse-bias voltage (as long as it remains well below the diode's breakdown voltage; see the next section). Plots of the 1N4148 silicon small-signal diode IV characteristic curves for both forward and reverse bias and at two temperatures are provided in the main text, Figure 2.

- ***Zener and avalanche breakdown***

As the reverse-bias voltage on a PN junction is increased, the intensity of the electric field in the depletion layer rises; it is particularly intense at the interface between the P- and N-type areas. Minority carriers entering the depletion layer are accelerated by the field; when their kinetic energies reach a few eV or so, collisions with atoms in the lattice may knock valence electrons out of them, creating additional electron-hole pairs. These newly-created charge carriers are also accelerated by the field and can create even more carriers as they collide with lattice atoms.

At sufficiently high reverse voltages this collision-induced ionization process may lead to an *avalanche* of additional charge carriers, and the reverse current will grow exponentially with increasing voltage beyond some reverse-bias threshold. This is the *avalanche breakdown* process, and the reverse-bias voltage threshold for its action is the diode's *reverse breakdown voltage*. The electric field intensity for any particular applied reverse-bias voltage depends on the impurity concentrations and the abruptness with which these concentrations change near the P-type and N-type interface, so a target reverse breakdown voltage may be engineered into a particular diode type.

Another effect of a very intense electric field in the depletion layer is the large electric polarization of the atoms in the lattice it induces: at field strengths $\gtrsim 10^6\,\mathrm{V/m}$ the potential difference across a distance of about 100Å can exceed the semiconductor's gap voltage. In this case a valence electron may quantum mechanically *tunnel* across this distance into the conduction band, creating an electron-hole pair; because of this tunneling process the electric field required to ionize a lattice atom is much smaller than it would need to be to ionize a single, independent atom ($\sim 10^{10}$–$10^{11}\,\mathrm{V/m}$); this effect was first theorized by the American physicist Clarence Zener in 1934. The tunneling rate grows exponentially as the required tunneling distance (inversely proportional to electric field strength) decreases, again leading to a large increase in reverse current (breakdown) as applied reverse-bias voltage exceeds the tunneling threshold.

Some diodes are purposely designed to be used in their reverse-bias breakdown region. These types of diodes are very useful as voltage references, simple voltage regulators, and overvoltage protection devices. The target reverse breakdown voltage may be engineered by adjusting a diode's impurity concentration and doping profile.

Diodes with reverse-breakdown voltages exceeding 6V or so are dominated by the avalanche breakdown process; those below 5V are predominantly subject to Zener breakdown. Regardless of breakdown voltage, those diodes designed to be used as voltage regulators with precisely-tailored reverse breakdown voltages are collectively called *Zener diodes*; those with high current-handling capacity and extremely fast response to voltages exceeding their breakdown threshold are usually called *avalanche diodes* and are primarily used for *transient voltage suppression* (TVS) and overvoltage protection.

Figure A-5 shows a typical Zener diode reverse-bias IV characteristic curve — a plot of the relationship between applied reverse-bias voltage and resulting current flow through the diode. The very steep portion of the curve corresponds to the diode's reverse-bias breakdown region; because the curve in this region is so steep, you can see that changes in the diode reverse current correspond to very small changes in reverse-bias voltage. Thus, the voltage across the Zener diode in this breakdown region is very insensitive to changes in the current through it. This characteristic makes the Zener diode useful as a simple voltage regulator.

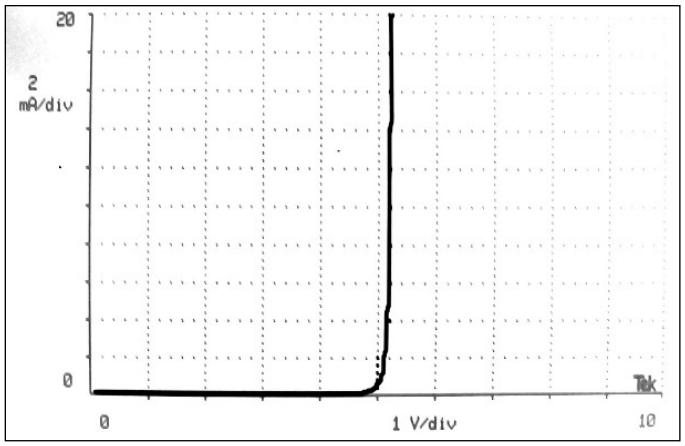

Figure A-5: Measured Zener diode I-V curve. The reverse diode current is plotted as a function of the applied reverse-bias voltage. As the applied voltage exceeds 5V, the diode current dramatically increases as the diode suffers reverse breakdown. As can be seen from the plot, diode reverse currents above about 15 mA correspond to a reverse-bias voltage of 5.3V.

- ***Thermal behavior of the charge carriers***

Finally, this section provides some brief statistical mechanical arguments and calculations to justify the assertions regarding the conduction electron and hole densities given in (13.A.1) and (13.A.2).You need not study this section unless you are unsure of the validity of those expressions. To follow the logic in this section it would be wise to review the sections concerning *fermions* and the *Fermi-Dirac distribution* in General Appendix B, ***Fundamental Concepts of Thermal Physics***. The following text refers to concepts and expressions from that discussion wherever it is convenient.

A pure semiconductor (no dopants) at $T = 0$ will have a full valence band and an empty conduction band. Even at room temperature it will be true that $E_g \gg k_B T$, and therefore the occupation probabilities of valence band single-electron states are very nearly 1 and those of the conduction band are very nearly 0. This result is consistent with the Fermi-Dirac distribution of equation (B.23) of General Appendix B only if the *chemical potential* μ is located in the energy gap relatively far from its edges — the energy of the top of the valence band (E_V) and the bottom of the conduction band (E_C). In fact, μ is usually very near the center of the band gap of a pure semiconductor. Thus the probability that a typical conduction-band single-electron state is occupied is

$$f_{\Delta E} = \frac{1}{e^{(E_C+\Delta E-\mu)/k_B T}+1} \approx e^{-(E_C+\Delta E-\mu)/k_B T} = e^{-(E_C-\mu)/k_B T}\, e^{-\Delta E/k_B T} \qquad (13.A.6)$$

where ΔE is the energy difference between the state and the bottom of the conduction band. Since the energy width of the conduction band is on the order of a few eV ($\gg k_B T$), the momentum-space structure of these states near the bottom of the conduction band is analogous to that of the free and independent electrons in a box considered in General Appendix B. Thus the conduction band density of single-electron states (within several $k_B T$ of E_C) is given by expression (B.17) to be

$$g(\Delta E) = 8\pi \frac{m^{3/2}}{h^3}\sqrt{2\Delta E}$$

and the expected number density of the electrons in the conduction band (n_c) is

$$n_c = N_C(T)\, e^{-(E_C-\mu)/k_B T}$$

$$\text{where:} \quad N_C(T) = \int_0^\infty g(\Delta E)\, e^{-\Delta E/k_B T}\, d(\Delta E) = \frac{(8\pi m_c\, k_B T)^{3/2}}{4h^3} \qquad (13.A.7)$$

You may think of N_C as the "effective" number of single-electron states (per volume) available in the conduction band (within a few $k_B T$ of the band edge). A similar

calculation for the probability that a typical single-electron state near the top of the valence band is empty $(1-f)$ and the resulting number density of holes in the valence band results in (13.A.8).

$$p_v = P_V(T)\, e^{-(\mu - E_V)/k_B T}$$
$$\text{where:} \quad P_V(T) = \frac{(8\pi m_v k_B T)^{3/2}}{4h^3} \tag{13.A.8}$$

It should be noted that m_c and m_v in these equations are the conduction electron and hole *effective masses* in the periodic potential of the semiconductor crystal; they are each within a factor of order unity of the free electron mass in many common semiconductor materials. Note that the conduction electron and hole densities given by (13.A.7) and (13.A.8) are valid even for doped semiconductors so long as the approximation in (13.A.6) is valid, i.e. the charge carriers are not *degenerate*. This will turn out to be the case so long as the dopant concentration is not too large and the semiconductor is not too much hotter than room temperature.

An expression for n_c and p_v which doesn't involve the chemical potential μ may be formed by taking the product of (13.A.7) and (13.A.8):

$$n_c p_v = N_C P_V\, e^{-(E_C - E_V)/k_B T} = N_C P_V\, e^{-E_g/k_B T} \tag{13.A.9}$$

This expression is an example of the *principle of mass action* or *detailed balance*: the right-hand side of equation (13.A.9) is proportional to the rate that electron-hole pairs will be thermally generated, whereas the left-hand side, the product of the electron and hole densities, is proportional to the rate that conduction electrons and holes will wander across one another and recombine. These two rates must balance when the system is in thermal equilibrium and the conduction electron and hole densities have become stable.

When the semiconductor is pure, then the only source of charge carriers is thermal generation from the intrinsic semiconductor atoms. In this case $n_c = p_v \equiv n_i$, the *intrinsic charge carrier density*. From equation (13.A.9) we immediately see that this implies that $n_i \propto T^{3/2} \exp[-E_g/(2k_B T)]$, as stated in (13.A.1). Comparing this result to either (13.A.7) or (13.A.8) shows that the chemical potential μ for a pure semiconductor must be near the center of the band gap, as stated earlier (its separation from the gap center is within a factor of order unity times $k_B T$). Note that since n_i^2 is given by the right-hand side of (13.A.9), but that the expression is correct even for impure (doped) semiconductors, then it must be the case that $n_c p_v = n_i^2$ even when a semiconductor is dominated by extrinsic charge carriers, as stated in (13.A.2).

When the introduction of dopant impurities drives the majority charge carrier density $\gg n_i$, the chemical potential μ will be repositioned toward the majority-carrier band gap edge. For example, the addition of 1 part per million phosphorous to a silicon crystal (introducing $\sim 10^{17}$ extrinsic conduction electrons per cm^3, about $10^8 n_i$) would reposition μ toward the conduction band by more than $18 k_B T$, or about 0.46 eV at room temperature (cf. equation (13.A.7) with $n_c \approx 10^8 n_i$). Since the silicon band gap energy is 1.12 eV, this would bring μ to about 0.1 eV below E_C. A similar concentration of an acceptor impurity such as aluminum would move μ to within about 0.1 eV of E_V. A PN junction formed from these two doped materials would have a contact potential determined by the difference in their chemical potentials (about 0.9 eV): close to, but less than, the semiconductor gap energy.