# PN junctions

This worksheet and all related files are licensed under the Creative Commons Attribution License, version 1.0. To view a copy of this license, visit http://creativecommons.org/licenses/by/1.0/, or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA. The terms and conditions of this license allow for free copying, distribution, and/or modification of all licensed works by the general public.

---

Resources and methods for learning about these subjects (list a few here, in preparation for your research):

Question 1

When "P" and "N" type semiconductor pieces are brought into close contact, free electrons from the "N" piece will rush over to fill holes in the "P" piece, creating a zone on both sides of the contact region devoid of charge carriers. What is this zone called, and what are its electrical characteristics?

file 00715

Answer 1

This is called the *depletion region*, and it is essentially an insulator at room temperatures.

Notes 1

Students should know that both "N" and "P" type semiconductors are electrically conductive. So, when a depletion region forms in the contact zone between two differing semiconductor types, the conductivity from end-to-end must be affected. Ask your students what this effect is, and what factors may influence it.

Question 2

What happens to the thickness of the depletion region in a PN junction when an external voltage is applied to it?
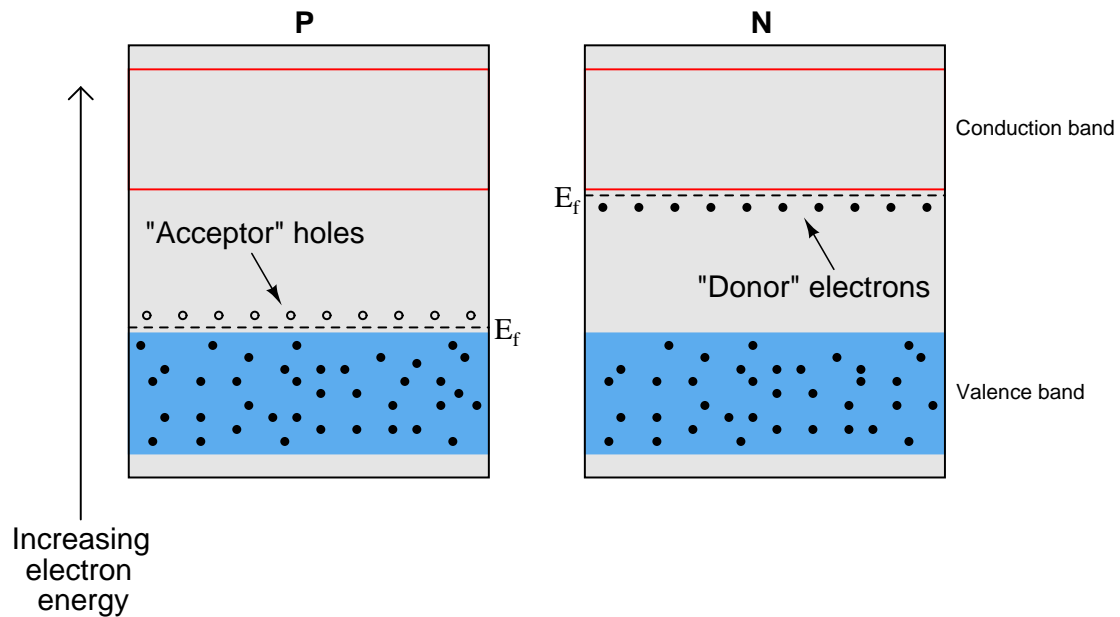
file 00909

Answer 2

The answer to this question depends entirely on the polarity of the applied voltage! One polarity tends to expand the depletion region, while the opposite polarity tends to compress it. I'll let you determine which polarity performs which action, based on your research.

Notes 2

Ask your students what effect this change in depletion layer thickness has on overall conductivity through the PN junction. Under what conditions will the conductivity be greatest, and under what conditions will the conductivity be least?

Question 3

Shown here are two energy diagrams: one for a "P" type semiconducting material and another for an "N" type.

**P**

"Acceptor" holes

$E_f$

Increasing electron energy

**N**

$E_f$

"Donor" electrons

Conduction band

Valence band

Next is an energy diagram showing the *initial* state when these two pieces of semiconducting material are brought into contact with each other. This is known as a *flatband diagram*:

*"Flatband" diagram -- temporary state!*

**P**

**N**

"Acceptor" holes

"Donor" electrons

$E_{f2}$

$E_{f1}$

Conduction band

Valence band
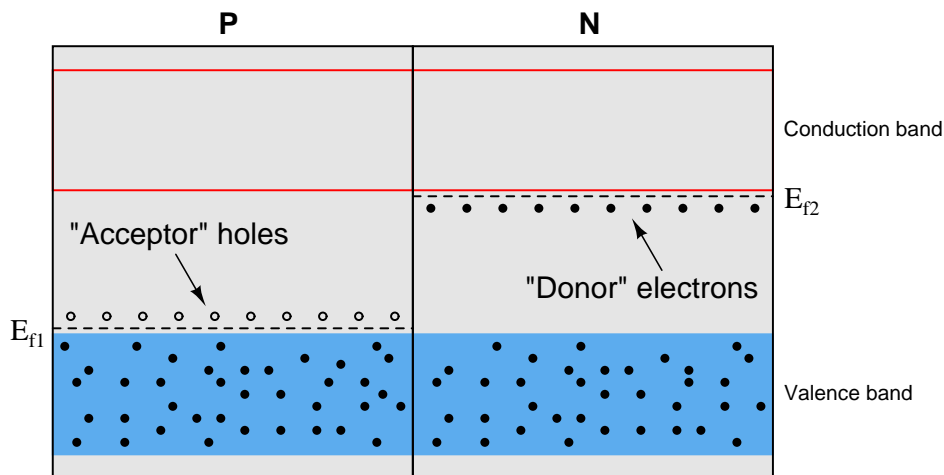
The state represented by the "flatband" diagram is most definitely a temporary one. The two different Fermi levels are incompatible with one another in the absence of an external electric field.

Draw a new energy diagram representing the final energy states after the two Fermi levels have equalized.

*Note: $E_f$ represents the Fermi energy level, and not a voltage. In physics, E always stands for energy and V for electric potential (voltage).*

file 00906

4

Answer 3
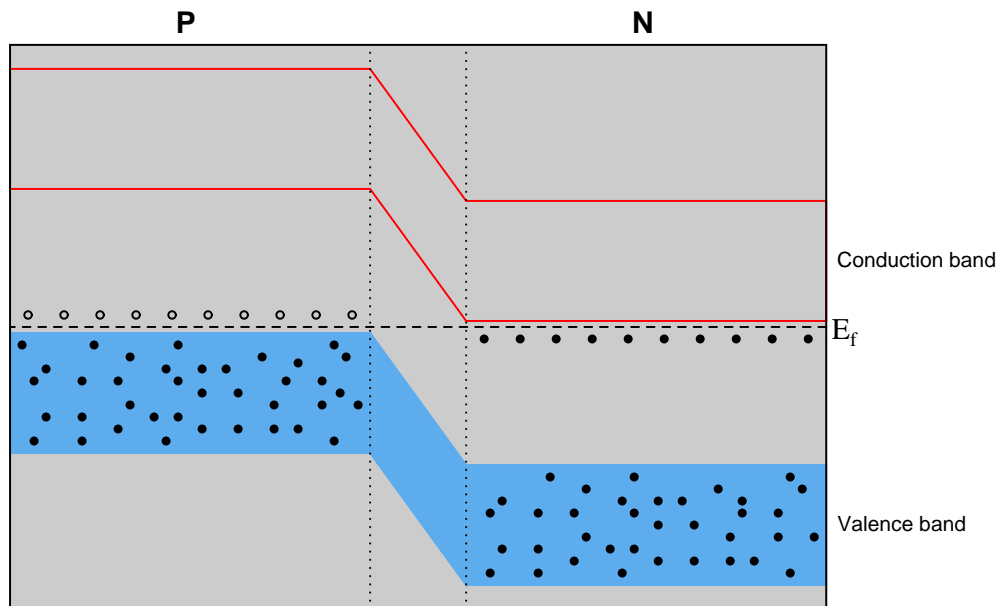


Electrons from the N-piece rushed over to fill holes in the P-piece *in order to achieve a lower energy state* and equalize the two Fermi levels. This displacement of charge carriers created an electric field which accounts for the sloped energy bands in the middle region.

Follow-up question: what is this middle region called?

Notes 3

This is one of those concepts I just couldn't understand when I had no comprehension of the quantum nature of electrons. In the "planetary" atomic model, there is no reason whatsoever for electrons to move from the N-piece to the P-piece unless there was an electric field pushing them in that direction. And conversely, once an electric field was created by the imbalance of electrons, the free-wheeling planetary theory would have predicted that the electrons move right back where they came from in order to neutralize the field.

Once you grasp the significance of quantized energy states, and the principle that particles do not "hold on" to unnecessary energy and therefore remain in high states when they could move down to a lower level, the concept becomes much clearer.

Question 4

Draw an energy diagram for a PN semiconductor junction under the influence of a *reverse* external voltage.

file 00910

Answer 4
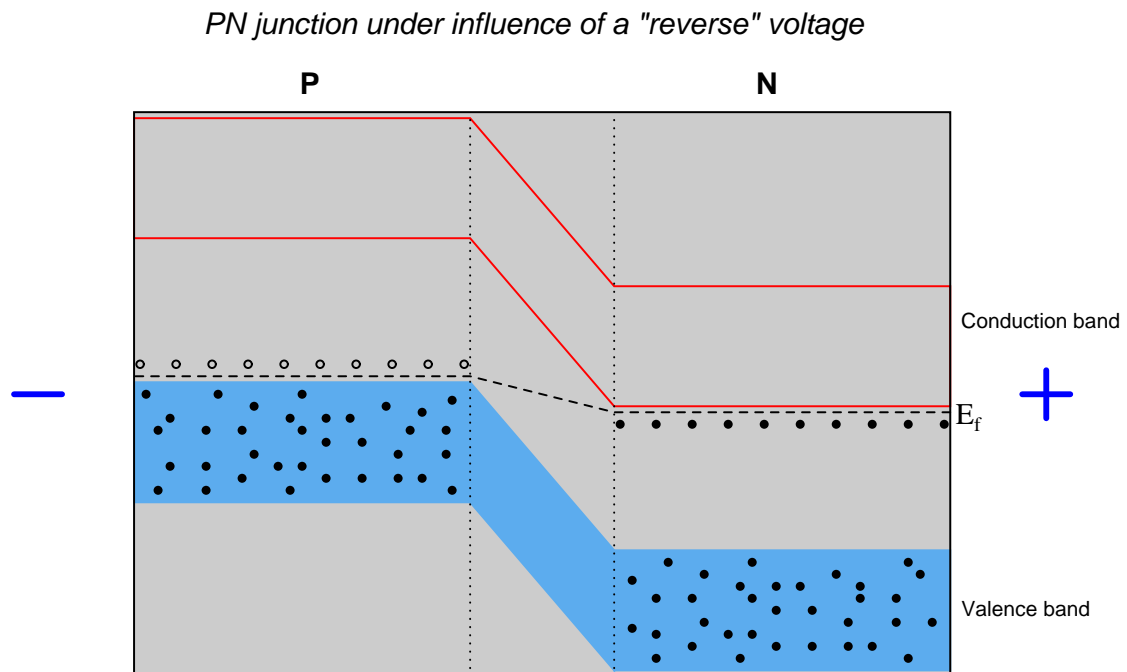
PN junction under influence of a "reverse" voltage



*Note: $E_f$ represents the Fermi energy level, and not a voltage. In physics, E always stands for energy and V for electric potential (voltage).*
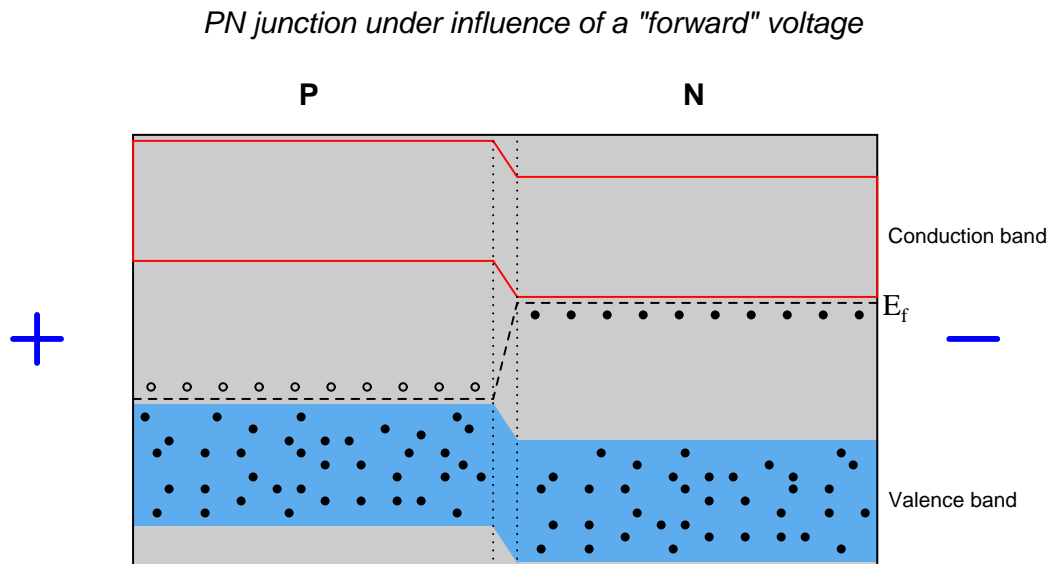
Notes 4

Here it is very important that students understand the effects an electric field has on energy bands.

Question 5

Draw an energy diagram for a PN semiconductor junction under the influence of a *forward* external voltage.

file 02005

Answer 5

*PN junction under influence of a "forward" voltage*



*Note: $E_f$ represents the Fermi energy level, and not a voltage. In physics, E always stands for energy and V for potential (voltage).*

Notes 5

Here it is very important that students understand the effects an electric field has on energy bands.
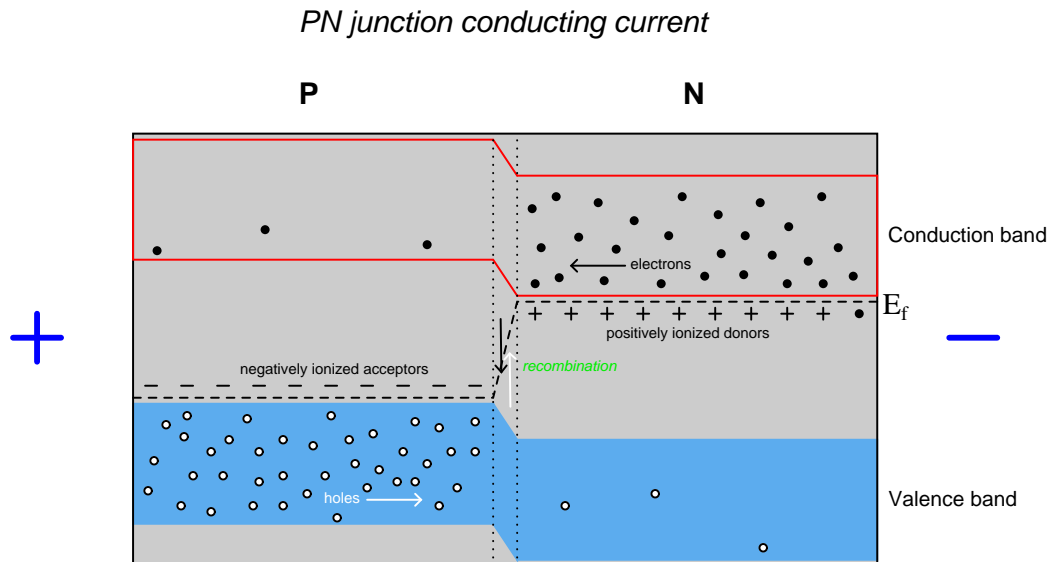
Question 6

Draw an energy diagram for a PN semiconductor junction showing the motion of electrons and holes conducting an electric current.

file 02006

Answer 6

For the sake of clearly seeing the actions of charge carriers (mobile electrons and holes), non-moving electrons in the valence bands are not shown:

*PN junction conducting current*



The "+" and "-" signs show the locations of ionized acceptor and donor atoms, having taken on electric charges to create valence-band holes and conduction-band electrons, respectively.

*Note: $E_f$ represents the Fermi energy level, and not a voltage. In physics, E always stands for energy and V for potential (voltage).*

Notes 6

Students will probably ask why there are a few holes shown in the N-type valence band, and why there are a few electrons in the P-type conduction band. Let them know that just because N-type materials are specifically designed to have conduction-band electrons does not mean they are completely devoid of valence-band holes, and visa-versa! What your students see here are *minority carriers*.

Question 7

The characteristically colored glow from a gas-discharge electric light is the result of energy emitted by electrons in the gas atoms as they fall from high-level "excited" states back to their natural ("ground") states. As a general rule of electron behavior, they must absorb energy from an external source to leap into a higher level, and they release that energy upon returning to their original level.

Given the existence of this phenomenon, what do you suspect might be occurring inside a PN junction as it conducts an electric current?

file 00911

Answer 7

PN junctions emit energy of a characteristic wavelength when conducting current. For some types of PN junctions, the wavelengths are within the visible range of light.

Follow-up question: what practical application can you think of for this phenomenon?

Notes 7

The practical application of this phenomenon should be obvious, and it is very commonplace in modern electronic equipment. Discuss with your students the energy-efficiency of this light emission as compared to an incandescent lamp.

Question 8

Is this diode *forward-biased* or *reverse biased?*
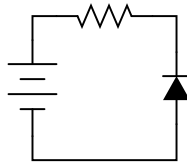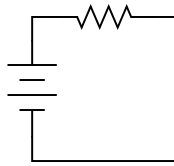


file 02299

Answer 8

This diode is reverse-biased.
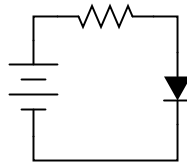
Notes 8

Nothing much to comment on here!

Question 9

Insert a diode into this circuit schematic in the correct direction to make it *forward-biased* by the battery voltage:



file 02300

Answer 9



Notes 9

Nothing much to comment on here!

Question 10

Most introductory textbooks will tell you that a silicon PN junction drops 0.7 volts when forward-biased, and a germanium PN junction drops 0.3 volts when forward biased. Design a circuit that tests the "forward voltage" ($V_F$) of a PN-junction diode, so you may measure the voltage yourself, without the use of a special diode-testing meter.

file 00711

Answer 10
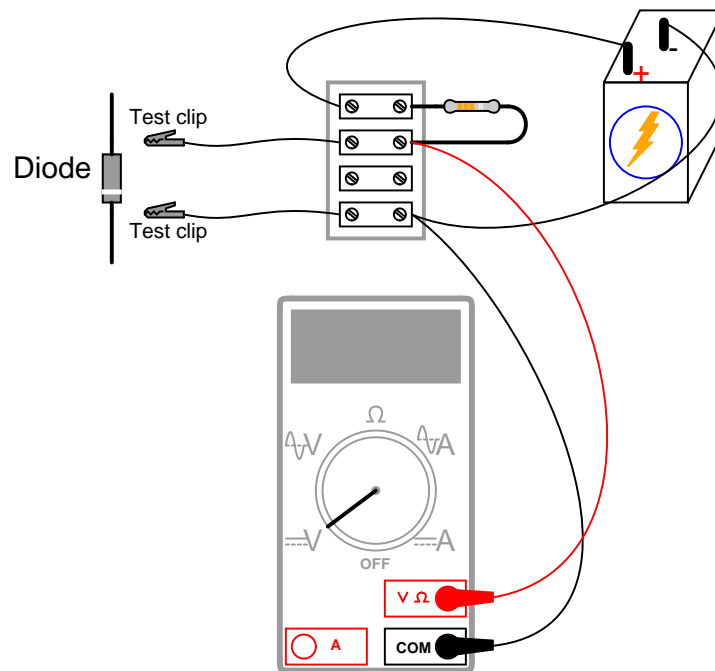


Notes 10

Ask your students how they would determine the size of resistor to use for this "diode test" circuit. Would be permissible to use any arbitrary value of resistor, or does the value matter significantly?

## Question 11

If a semiconductor PN junction is reverse-biased, ideally no continuous current will go through it. However, in real life there will be a small amount of reverse-bias current that goes through the junction. How is this possible? What allows this reverse current to flow?

file 00913

## Answer 11

Minority carriers allow reverse current through a PN junction.

## Notes 11

Review with your students what "minority carriers" are, and apply this concept to the PN junction. Trace the motions of these minority carriers, and compare them with the motions of majority carriers in a forward-biased PN junction.

Question 12

The relationship between voltage and current for a PN junction is described by this equation, sometimes referred to as the "diode equation," or "Shockley's diode equation" after its discoverer:

$$I_D = I_S(e^{\frac{qV_D}{NkT}} - 1)$$

Where,

$I_D$ = Current through the PN junction, in amps

$I_S$ = PN junction saturation current, in amps (typically 1 picoamp)

$e$ = Euler's number $\approx 2.718281828$

$q$ = Electron unit charge, $1.6 \times 10^{-19}$ coulombs

$V_D$ = Voltage across the PN junction, in volts

$N$ = Nonideality coefficient, or emission coefficient (typically between 1 and 2)

$k$ = Boltzmann's constant, $1.38 \times 10^{-23}$

$T$ = Junction temperature, degrees Kelvin

At first this equation may seem very daunting, until you realize that there are really only three variables in it: $I_D$, $V_D$, and $T$. All the other terms are constants. Since in most cases we assume temperature is fairly constant as well, we are really only dealing with two variables: diode current and diode voltage. Based on this realization, re-write the equation as a proportionality rather than an equality, showing how the two variables of diode current and voltage relate:

$$I_D \propto \; . \; . \; .$$

Based on this simplified equation, what would an I/V graph for a PN junction look like? How does this graph compare against the I/V graph for a resistor?
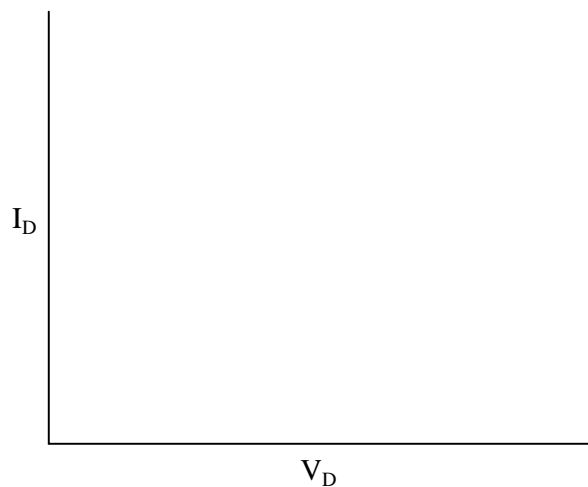


file 00712

Answer 12

Simplified proportionality:

$$I_D \propto e^{V_D}$$

The graph described by the "diode formula" is a standard exponential curve, rising sharply as the independent variable ($V_D$, in this case) increases. The corresponding graph for a resistor, of course, is linear.

14

Notes 12

Ask your students to sketch their own renditions of an exponential curve on the whiteboard for all to see. Don't just let them get away with parroting the answer: "It's an exponential curve."

## Question 13

Shockley's diode equation in standard form is quite lengthy, but it may be considerably simplified for conditions of room temperature. Note that if the temperature ($T$) is assumed to be room temperature ($25^o$ C), there are three constants in the equation that are the same for all PN junctions: $T$, $k$, and $q$.

$$I_D = I_S(e^{\frac{qV_D}{NkT}} - 1)$$

The quantity $\frac{kT}{q}$ is known as the *thermal voltage* of the junction. Calculate the value of this thermal voltage, given a room temperature of $25^o$ C. Then, substitute this quantity into the original "diode formula" so as to simplify its appearance.

file 00713

## Answer 13

If you obtained an answer of 2.16 mV for the "thermal voltage," you have the temperature figure in the wrong units!

$$I_D = I_S(e^{\frac{V_D}{0.0257N}} - 1)$$

## Notes 13

Of course, students will have to research the difference between degrees Kelvin and degrees Celsius to successfully calculate the thermal voltage for the junction. They will also have to figure out how to substitute this figure in place of $q$, $k$, and $T$ in the original equation. The latter step will be difficult for students not strong in algebra skills.

For those students, I would suggest posing the following question to get them thinking properly about algebraic substitution. Suppose we had the formula $y = x^{\frac{ab}{cd}}$, and we knew that $\frac{b}{c}$ could be written as $m$. How would we substitute $m$ into the original equation? Answer: $y = x^{\frac{am}{d}}$.

Question 14

A student sets up a circuit that looks like this, to gather data for characterizing a diode:
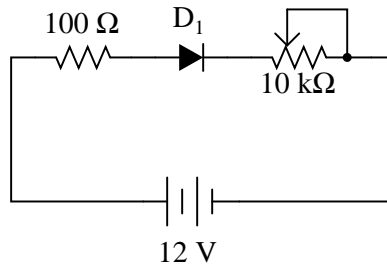


Measuring diode voltage and diode current in this circuit, the student generates the following table of data:

| $V_{diode}$ | $I_{diode}$ |
|-------------|-------------|
| 0.600 V | 1.68 mA |
| 0.625 V | 2.88 mA |
| 0.650 V | 5.00 mA |
| 0.675 V | 8.68 mA |
| 0.700 V | 14.75 mA |
| 0.725 V | 27.25 mA |
| 0.750 V | 48.2 mA |

This student knows that the behavior of a PN junction follows Shockley's diode equation, and that the equation may be simplified to the following form:

$$I_{diode} = I_S(e^{\frac{V_{diode}}{K}} - 1)$$

Where,
$K$ = a constant incorporating both the thermal voltage and the nonideality coefficient

The goal of this experiment is to calculate $K$ and $I_S$, so that the diode's current may be predicted for any arbitrary value of voltage drop. However, the equation must be simplified a bit before the student can proceed.

At substantial levels of current, the exponential term is very much larger than unity ($e^{\frac{V_{diode}}{K}} >> 1$), so the equation may be simplified as such:

$$I_{diode} \approx I_S(e^{\frac{V_{diode}}{K}})$$

From this equation, determine how the student would calculate $K$ and $I_S$ from the data shown in the table. Also, explain how this student may verify the accuracy of these calculated values.
<u>file 01923</u>

Answer 14

$K \approx 0.04516$

$I_S \approx 2.869$ nA

Hint: this may be a difficult problem to solve if you are unfamiliar with the algebraic technique of dividing one equation by another. Here is the technique shown in general terms:

$$\text{Given:} \quad y_1 = ax_1 \qquad y_2 = ax_2$$

$$\frac{y_1}{y_2} = \frac{ax_1}{ax_2}$$

From here, it may be possible to perform simplifications impossible before. I suggest using this technique to solve for $K$ first.

Follow-up question: explain how this student knew it was "safe" to simplify the Shockley diode equation by eliminating the "- 1" term. Is this sort of elimination always permissible? Why or why not?

Notes 14

The algebraic technique used to solve for $K$ is very useful for certain types of problems.

Discuss the follow-up question with your students. It is important in the realm of technical mathematics to have a good sense of the relative values of equation terms, so that one may "safely" eliminate terms as a simplifying technique without incurring significant errors. In the Shockley diode equation it is easy to show that the exponential term is *enormous* compared to 1 for the values of $V_{diode}$ shown in the table (assuming a typical value for thermal voltage), and so the "- 1" part is very safe to eliminate.

Also discuss the idea of verifying the calculated values of $K$ and $I_S$ with your students, to help them cultivate a scientifically critical point of view in their study of electronics.

Incidentally, the data in this table came from a real experiment, set up exactly as shown by the schematic diagram in the question. Care was taken to avoid diode heating by turning the potentiometer to maximum resistance between readings.

Question 15

In order to simplify analysis of circuits containing PN junctions, a "standard" forward voltage drop is assumed for any conducting junction, the exact figure depending on the type of semiconductor material the junction is made of.

How much voltage is assumed to be dropped across a conducting *silicon* PN junction? How much voltage is assumed for a forward-biased *germanium* PN junction? Identify some factors that cause the real forward voltage drop of a PN junction to deviate from its "standard" figure.

file 00898

Answer 15

Silicon = 0.7 volts ; Germanium = 0.3 volts.

Temperature, current, and doping concentration all affect the forward voltage drop of a PN junction.

Notes 15

I've seen too many students gain the false impression that silicon PN junctions *always* drop 0.7 volts, no matter what the conditions. This "fact" is emphasized so strongly in many textbooks that students usually don't think to ask when they measure a diode's forward voltage drop and find it to be considerably different than 0.7 volts! It is very important that students realize this figure is an approximation only, used for the sake of (greatly) simplifying junction semiconductor circuit analysis.

Question 16

$\boxed{\int f(x)\,dx \quad \textit{Calculus alert!}}$

A forward-biased PN semiconductor junction does not possess a "resistance" in the same manner as a resistor or a length of wire. Any attempt at applying Ohm's Law to a diode, then, is doomed from the start.

This is not to say that we cannot assign a *dynamic* value of resistance to a PN junction, though. The fundamental definition of resistance comes from Ohm's Law, and it is expressed in derivative form as such:

$$R = \frac{dV}{dI}$$

The fundamental equation relating current and voltage together for a PN junction is Shockley's diode equation:

$$I = I_S(e^{\frac{qV}{NkT}} - 1)$$

At room temperature (approximately 21 degrees C, or 294 degrees K), the thermal voltage of a PN junction is about 25 millivolts. Substituting 1 for the nonideality coefficient, we may simply the diode equation as such:

$$I = I_S(e^{\frac{V}{0.025}} - 1) \qquad \text{or} \qquad I = I_S(e^{40V} - 1)$$

Differentiate this equation with respect to $V$, so as to determine $\frac{dI}{dV}$, and then reciprocate to find a mathematical definition for dynamic resistance ($\frac{dV}{dI}$) of a PN junction. Hints: saturation current ($I_S$) is a very small constant for most diodes, and the final equation should express dynamic resistance in terms of thermal voltage (25 mV) and diode current ($I$).

file 02538

Answer 16

$$r \approx \frac{25 \text{ mV}}{I}$$

The result of this derivation is important in the analysis of certain transistor amplifiers, where the dynamic resistance of the base-emitter PN junction is significant to bias and gain approximations. I show the solution steps for you here because it is a neat application of differentiation (and substitution) to solve a real-world problem:

$$I = I_S(e^{40V} - 1)$$

$$\frac{dI}{dV} = I_S(40e^{40V} - 0)$$

$$\frac{dI}{dV} = 40I_Se^{40V}$$

Now, we manipulate the original equation to obtain a definition for $I_Se^{40V}$ in terms of current, for the sake of substitution:

$$I = I_S(e^{40V} - 1)$$

$$I = I_Se^{40V} - I_S$$

$$I + I_S = I_Se^{40V}$$

Substituting this expression into the derivative:

$$\frac{dI}{dV} = 40(I + I_S)$$

Reciprocating to get voltage over current (the proper form for resistance):

$$\frac{dV}{dI} = \frac{0.025}{I + I_S}$$

Now we may get rid of the saturation current term, because it is negligibly small:

$$\frac{dV}{dI} \approx \frac{0.025}{I}$$

$$r \approx \frac{25 \text{ mV}}{I}$$

The constant of 25 millivolts is not set in stone, by any means. Its value varies with temperature, and is sometimes given as 26 millivolts or even 30 millivolts.

## Question 17

Measure the forward voltage drop of a silicon rectifying diode, such as a model 1N4001. How close is the measured forward voltage drop to the "ideal" figure usually assumed for silicon PN junctions? What happens when you increase the temperature of the diode by holding on to it with your fingers? What happens when you decrease the temperature of the diode by touching an ice cube to it?

file 00714

## Answer 17

Did you really think I was going to give away the answer here, and spoil the fun of setting up an experiment?

## Notes 17

Diodes are quite temperature-sensitive, so this experiment will be very easy to conduct. You may not have ice available in your classroom, but that's okay. Your students should realize that experiments such as this are perfectly fair to perform at home, where they probably do have access to ice.

Question 18

The nonconducting depletion region of a PN junction forms a parasitic capacitance between the P and the N semiconductor region. Does the capacitance increase or decrease as a greater reverse-bias voltage is applied to the PN junction? Explain your answer.

file 00899

Answer 18

The junction capacitance will decrease as the reverse-bias voltage across the junction increases.

Challenge question: can you think of any practical applications for this variable-capacitance effect?

Notes 18

This question is a good review of capacitor theory, and also an opportunity to introduce a special kind of diode: the *varactor*.

**Don't just sit there! Build something!!**

Learning to mathematically analyze circuits requires much study and practice. Typically, students practice by working through lots of sample problems and checking their answers against those provided by the textbook or the instructor. While this is good, there is a much better way.

You will learn much more by actually *building and analyzing real circuits*, letting your test equipment provide the "answers" instead of a book or another person. For successful circuit-building exercises, follow these steps:

1. Carefully measure and record all component values prior to circuit construction, choosing resistor values high enough to make damage to any active components unlikely.
2. Draw the schematic diagram for the circuit to be analyzed.
3. Carefully build this circuit on a breadboard or other convenient medium.
4. Check the accuracy of the circuit's construction, following each wire to each connection point, and verifying these elements one-by-one on the diagram.
5. Mathematically analyze the circuit, solving for all voltage and current values.
6. Carefully measure all voltages and currents, to verify the accuracy of your analysis.
7. If there are any substantial errors (greater than a few percent), carefully check your circuit's construction against the diagram, then carefully re-calculate the values and re-measure.

When students are first learning about semiconductor devices, and are most likely to damage them by making improper connections in their circuits, I recommend they experiment with large, high-wattage components (1N4001 rectifying diodes, TO-220 or TO-3 case power transistors, etc.), and using dry-cell battery power sources rather than a benchtop power supply. This decreases the likelihood of component damage.

As usual, avoid very high and very low resistor values, to avoid measurement errors caused by meter "loading" (on the high end) and to avoid transistor burnout (on the low end). I recommend resistors between 1 kΩ and 100 kΩ.

One way you can save time and reduce the possibility of error is to begin with a very simple circuit and incrementally add components to increase its complexity after each analysis, rather than building a whole new circuit for each practice problem. Another time-saving technique is to re-use the same components in a variety of different circuit configurations. This way, you won't have to measure any component's value more than once.

file 00505

---

Answer 19

Let the electrons themselves give you the answers to your own "practice problems"!

It has been my experience that students require much practice with circuit analysis to become proficient. To this end, instructors usually provide their students with lots of practice problems to work through, and provide answers for students to check their work against. While this approach makes students proficient in circuit theory, it fails to fully educate them.

Students don't just need mathematical practice. They also need real, hands-on practice building circuits and using test equipment. So, I suggest the following alternative approach: students should *build* their own "practice problems" with real components, and try to mathematically predict the various voltage and current values. This way, the mathematical theory "comes alive," and students gain practical proficiency they wouldn't gain merely by solving equations.

Another reason for following this method of practice is to teach students *scientific method*: the process of testing a hypothesis (in this case, mathematical predictions) by performing a real experiment. Students will also develop real troubleshooting skills as they occasionally make circuit construction errors.

Spend a few moments of time with your class to review some of the "rules" for building circuits before they begin. Discuss these issues with your students in the same Socratic manner you would normally discuss the worksheet questions, rather than simply telling them what they should and should not do. I never cease to be amazed at how poorly students grasp instructions when presented in a typical lecture (instructor monologue) format!

A note to those instructors who may complain about the "wasted" time required to have students build real circuits instead of just mathematically analyzing theoretical circuits:

*What is the purpose of students taking your course?*

If your students will be working with real circuits, then they should learn on real circuits whenever possible. If your goal is to educate theoretical physicists, then stick with abstract analysis, by all means! But most of us plan for our students to do something in the real world with the education we give them. The "wasted" time spent building real circuits will pay huge dividends when it comes time for them to apply their knowledge to practical problems.

Furthermore, having students build their own practice problems teaches them how to perform *primary research*, thus empowering them to continue their electrical/electronics education autonomously.

In most sciences, realistic experiments are much more difficult and expensive to set up than electrical circuits. Nuclear physics, biology, geology, and chemistry professors would just love to be able to have their students apply advanced mathematics to real experiments posing no safety hazard and costing less than a textbook. They can't, but you can. Exploit the convenience inherent to your science, and *get those students of yours practicing their math on lots of real circuits!*