



PEARSON NEW INTERNATIONAL EDITION

Electronic Devices and Circuit Theory
Robert L. Boylestad Louis Nashelsky
Eleventh Edition

Pearson New International Edition

Electronic Devices and Circuit Theory
Robert L. Boylestad Louis Nashelsky
Eleventh Edition

PEARSON

Pearson Education Limited

Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

© Pearson Education Limited 2014

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a licence permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

PEARSON

ISBN 10: 1-292-02563-8
ISBN 13: 978-1-292-02563-6

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

1. Semiconductor Diodes Robert L. Boylestad/Louis Nashelsky	1
2. Diode Applications Robert L. Boylestad/Louis Nashelsky	57
3. Bipolar Junction Transistors Robert L. Boylestad/Louis Nashelsky	133
4. DC Biasing—BJTs Robert L. Boylestad/Louis Nashelsky	165
5. BJT AC Analysis Robert L. Boylestad/Louis Nashelsky	261
6. Field-Effect Transistors Robert L. Boylestad/Louis Nashelsky	387
7. FET Biasing Robert L. Boylestad/Louis Nashelsky	433
8. FET Amplifiers Robert L. Boylestad/Louis Nashelsky	495
9. BJT and JFET Frequency Response Robert L. Boylestad/Louis Nashelsky	561
10. Operational Amplifiers Robert L. Boylestad/Louis Nashelsky	625
11. Op-Amp Applications Robert L. Boylestad/Louis Nashelsky	673
12. Power Amplifiers Robert L. Boylestad/Louis Nashelsky	705
13. Linear-Digital ICs Robert L. Boylestad/Louis Nashelsky	745

14. Feedback and Oscillator Circuits	
Robert L. Boylestad/Louis Nashelsky	775
15. Power Supplies (Voltage Regulators)	
Robert L. Boylestad/Louis Nashelsky	807
16. Other Two-Terminal Devices	
Robert L. Boylestad/Louis Nashelsky	837
17. pnpn and Other Devices	
Robert L. Boylestad/Louis Nashelsky	869
18. Appendix: Ripple Factor and Voltage Calculations	
Robert L. Boylestad/Louis Nashelsky	907
Index	913

Semiconductor Diodes

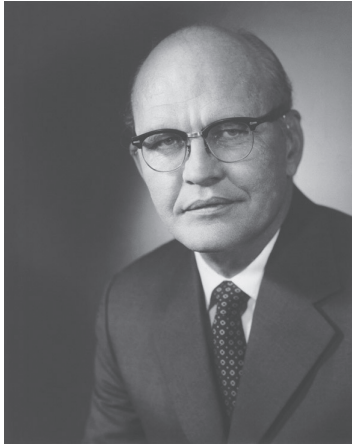
CHAPTER OBJECTIVES

- Become aware of the general characteristics of three important semiconductor materials: Si, Ge, GaAs.
- Understand conduction using electron and hole theory.
- Be able to describe the difference between n - and p -type materials.
- Develop a clear understanding of the basic operation and characteristics of a diode in the no-bias, forward-bias, and reverse-bias regions.
- Be able to calculate the dc, ac, and average ac resistance of a diode from the characteristics.
- Understand the impact of an equivalent circuit whether it is ideal or practical.
- Become familiar with the operation and characteristics of a Zener diode and light-emitting diode.

1 INTRODUCTION

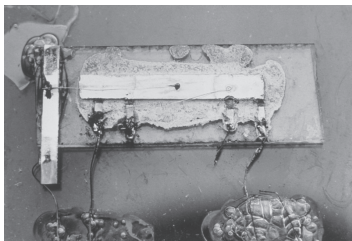
One of the noteworthy things about this field, as in many other areas of technology, is how little the fundamental principles change over time. Systems are incredibly smaller, current speeds of operation are truly remarkable, and new gadgets surface every day, leaving us to wonder where technology is taking us. However, if we take a moment to consider that the majority of all the devices in use were invented decades ago and that design techniques appearing in texts as far back as the 1930s are still in use, we realize that most of what we see is primarily a steady improvement in construction techniques, general characteristics, and application techniques rather than the development of new elements and fundamentally new designs. The result is that most of the devices discussed in this text have been around for some time, and that texts on the subject written a decade ago are still good references with content that has not changed very much. The major changes have been in the understanding of how these devices work and their full range of capabilities, and in improved methods of teaching the fundamentals associated with them. The benefit of all this to the new student of the subject is that the material in this text will, we hope, have reached a level where it is relatively easy to grasp and the information will have application for years to come.

The miniaturization that has occurred in recent years leaves us to wonder about its limits. Complete systems now appear on wafers thousands of times smaller than the single element of earlier networks. The first integrated circuit (IC) was developed by Jack Kilby while working at Texas Instruments in 1958 (Fig. 1). Today, the Intel® Core™ i7 Extreme



Jack St. Clair Kilby, inventor of the integrated circuit and co-inventor of the electronic handheld calculator. (Courtesy of Texas Instruments.)

Born: Jefferson City, Missouri, 1923.
MS, University of Wisconsin.
Director of Engineering and Technology, Components Group, Texas Instruments. Fellow of the IEEE.
Holds more than 60 U.S. patents.



The first integrated circuit, a phase-shift oscillator, invented by Jack S. Kilby in 1958. (Courtesy of Texas Instruments.)

FIG. 1

Jack St. Clair Kilby.

Edition Processor of Fig. 2 has 731 million transistors in a package that is only slightly larger than a 1.67 sq. inches. In 1965, Dr. Gordon E. Moore presented a paper predicting that the transistor count in a single IC chip would double every two years. Now, more than 45 years, later we find that his prediction is amazingly accurate and expected to continue for the next few decades. We have obviously reached a point where the primary purpose of the container is simply to provide some means for handling the device or system and to provide a mechanism for attachment to the remainder of the network. Further miniaturization appears to be limited by four factors: the quality of the semiconductor material, the network design technique, the limits of the manufacturing and processing equipment, and the strength of the innovative spirit in the semiconductor industry.

The first device to be introduced here is the simplest of all electronic devices, yet has a range of applications that seems endless.

2 SEMICONDUCTOR MATERIALS: Ge, Si, AND GaAs

The construction of every discrete (individual) solid-state (hard crystal structure) electronic device or integrated circuit begins with a semiconductor material of the highest quality.

Semiconductors are a special class of elements having a conductivity between that of a good conductor and that of an insulator.

In general, semiconductor materials fall into one of two classes: *single-crystal* and *compound*. Single-crystal semiconductors such as germanium (Ge) and silicon (Si) have a repetitive crystal structure, whereas compound semiconductors such as gallium arsenide (GaAs), cadmium sulfide (CdS), gallium nitride (GaN), and gallium arsenide phosphide (GaAsP) are constructed of two or more semiconductor materials of different atomic structures.

The three semiconductors used most frequently in the construction of electronic devices are Ge, Si, and GaAs.

In the first few decades following the discovery of the diode in 1939 and the transistor in 1947 germanium was used almost exclusively because it was relatively easy to find and was available in fairly large quantities. It was also relatively easy to refine to obtain very high levels of purity, an important aspect in the fabrication process. However, it was discovered in the early years that diodes and transistors constructed using germanium as the base material suffered from low levels of reliability due primarily to its sensitivity to changes in temperature. At the time, scientists were aware that another material, silicon, had improved temperature sensitivities, but the refining process for manufacturing silicon of very high levels of purity was still in the development stages. Finally, however, in 1954 the first silicon transistor was introduced, and silicon quickly became the semiconductor material of choice. Not only is silicon less temperature sensitive, but it is one of the most abundant materials on earth, removing any concerns about availability. The flood gates now opened to this new material, and the manufacturing and design technology improved steadily through the following years to the current high level of sophistication.

As time moved on, however, the field of electronics became increasingly sensitive to issues of speed. Computers were operating at higher and higher speeds, and communication systems were operating at higher levels of performance. A semiconductor material capable of meeting these new needs had to be found. The result was the development of the first GaAs transistor in the early 1970s. This new transistor had speeds of operation up to five times that of Si. The problem, however, was that because of the years of intense design efforts and manufacturing improvements using Si, Si transistor networks for most applications were cheaper to manufacture and had the advantage of highly efficient design strategies. GaAs was more difficult to manufacture at high levels of purity, was more expensive, and had little design support in the early years of development. However, in time the demand for increased speed resulted in more funding for GaAs research, to the point that today it is often used as the base material for new high-speed, very large scale integrated (VLSI) circuit designs.

This brief review of the history of semiconductor materials is not meant to imply that GaAs will soon be the only material appropriate for solid-state construction. Germanium devices are still being manufactured, although for a limited range of applications. Even though it is a temperature-sensitive semiconductor, it does have characteristics that find application in a limited number of areas. Given its availability and low manufacturing costs, it will continue to find its place in product catalogs. As noted earlier, Si has the benefit of years of development, and is the leading semiconductor material for electronic components and ICs. In fact, Si is still the fundamental building block for Intel's new line of processors.

3 COVALENT BONDING AND INTRINSIC MATERIALS

To fully appreciate why Si, Ge, and GaAs are the semiconductors of choice for the electronics industry requires some understanding of the atomic structure of each and how the atoms are bound together to form a crystalline structure. The fundamental components of an atom are the electron, proton, and neutron. In the lattice structure, neutrons and protons form the nucleus and electrons appear in fixed orbits around the nucleus. The Bohr model for the three materials is provided in Fig. 3.

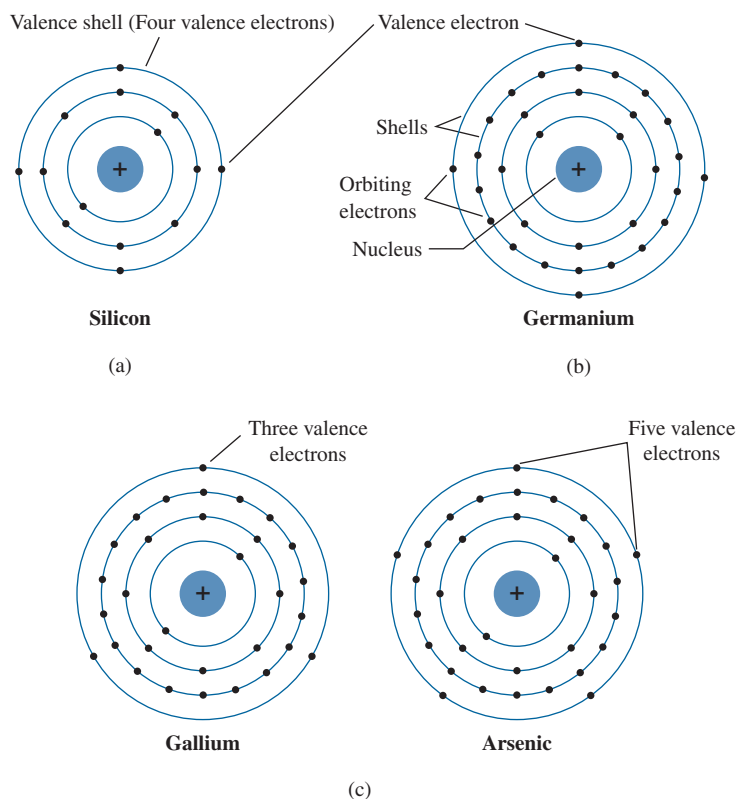


FIG. 3

Atomic structure of (a) silicon; (b) germanium; and (c) gallium and arsenic.

As indicated in Fig. 3, silicon has 14 orbiting electrons, germanium has 32 electrons, gallium has 31 electrons, and arsenic has 33 orbiting electrons (the same arsenic that is a very poisonous chemical agent). For germanium and silicon there are four electrons in the outermost shell, which are referred to as *valence electrons*. Gallium has three valence electrons and arsenic has five valence electrons. Atoms that have four valence electrons are called *tetravalent*, those with three are called *trivalent*, and those with five are called *pentavalent*. The term *valence* is used to indicate that the potential (ionization potential) required to remove any one of these electrons from the atomic structure is significantly lower than that required for any other electron in the structure.

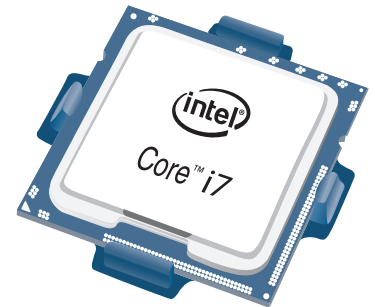
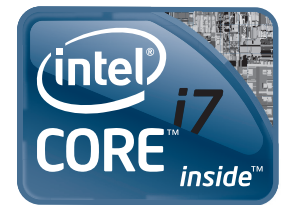
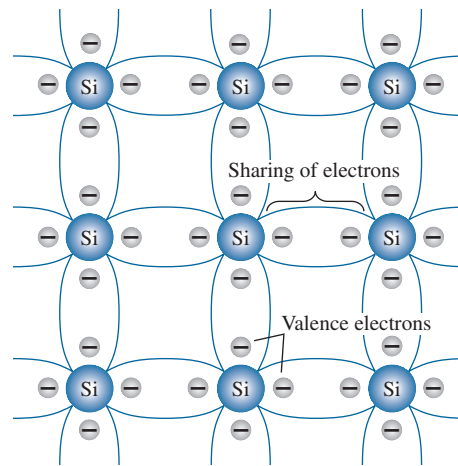


FIG. 2

Intel® Core™ i7 Extreme Edition Processor.

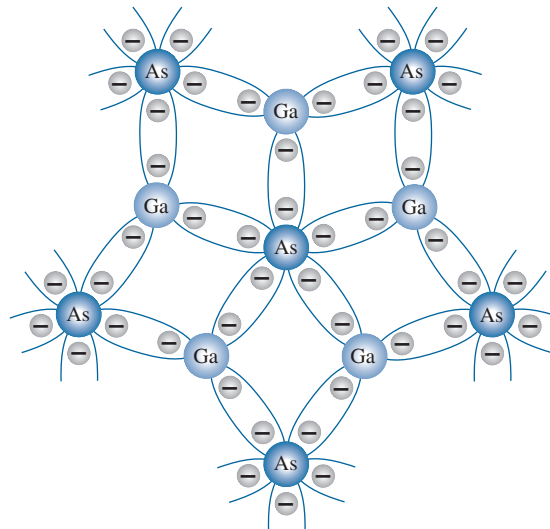
**FIG. 4**

Covalent bonding of the silicon atom.

In a pure silicon or germanium crystal the four valence electrons of one atom form a bonding arrangement with four adjoining atoms, as shown in Fig. 4.

This bonding of atoms, strengthened by the sharing of electrons, is called covalent bonding.

Because GaAs is a compound semiconductor, there is sharing between the two different atoms, as shown in Fig. 5. Each atom, gallium or arsenic, is surrounded by atoms of the complementary type. There is still a sharing of electrons similar in structure to that of Ge and Si, but now five electrons are provided by the As atom and three by the Ga atom.

**FIG. 5**

Covalent bonding of the GaAs crystal.

Although the covalent bond will result in a stronger bond between the valence electrons and their parent atom, it is still possible for the valence electrons to absorb sufficient kinetic energy from external natural causes to break the covalent bond and assume the “free” state. The term *free* is applied to any electron that has separated from the fixed lattice structure and is very sensitive to any applied electric fields such as established by voltage sources or any difference in potential. *The external causes include effects such as light energy in the form of photons and thermal energy (heat) from the surrounding medium.* At room temperature there are approximately 1.5×10^{10} free carriers in 1 cm^3 of *intrinsic* silicon material, that is, 15,000,000,000 (15 billion) electrons in a space smaller than a small sugar cube—an enormous number.

The term intrinsic is applied to any semiconductor material that has been carefully refined to reduce the number of impurities to a very low level—essentially as pure as can be made available through modern technology.

The free electrons in a material due only to external causes are referred to as *intrinsic carriers*. Table 1 compares the number of intrinsic carriers per cubic centimeter (abbreviated n_i) for Ge, Si, and GaAs. It is interesting to note that Ge has the highest number and GaAs the lowest. In fact, Ge has more than twice the number as GaAs. The number of carriers in the intrinsic form is important, but other characteristics of the material are more significant in determining its use in the field. One such factor is the *relative mobility* (μ_n) of the free carriers in the material, that is, the ability of the free carriers to move throughout the material. Table 2 clearly reveals that the free carriers in GaAs have more than five times the mobility of free carriers in Si, a factor that results in response times using GaAs electronic devices that can be up to five times those of the same devices made from Si. Note also that free carriers in Ge have more than twice the mobility of electrons in Si, a factor that results in the continued use of Ge in high-speed radio frequency applications.

TABLE 1
Intrinsic Carriers n_i

Semiconductor	Intrinsic Carriers (per cubic centimeter)
GaAs	1.7×10^6
Si	1.5×10^{10}
Ge	2.5×10^{13}

TABLE 2
Relative Mobility Factor μ_n

Semiconductor	μ_n ($\text{cm}^2/\text{V}\cdot\text{s}$)
Si	1500
Ge	3900
GaAs	8500

One of the most important technological advances of recent decades has been the ability to produce semiconductor materials of very high purity. Recall that this was one of the problems encountered in the early use of silicon—it was easier to produce germanium of the required purity levels. Impurity levels of 1 part in 10 billion are common today, with higher levels attainable for large-scale integrated circuits. One might ask whether these extremely high levels of purity are necessary. They certainly are if one considers that the addition of one part of impurity (of the proper type) per million in a wafer of silicon material can change that material from a relatively poor conductor to a good conductor of electricity. We obviously have to deal with a whole new level of comparison when we deal with the semiconductor medium. The ability to change the characteristics of a material through this process is called *doping*, something that germanium, silicon, and gallium arsenide readily and easily accept. The doping process is discussed in detail in Sections 5 and 6.

One important and interesting difference between semiconductors and conductors is their reaction to the application of heat. For conductors, the resistance increases with an increase in heat. This is because the numbers of carriers in a conductor do not increase significantly with temperature, but their vibration pattern about a relatively fixed location makes it increasingly difficult for a sustained flow of carriers through the material. Materials that react in this manner are said to have a *positive temperature coefficient*. Semiconductor materials, however, exhibit an increased level of conductivity with the application of heat. As the temperature rises, an increasing number of valence electrons absorb sufficient thermal energy to break the covalent bond and to contribute to the number of free carriers. Therefore:

Semiconductor materials have a negative temperature coefficient.

4 ENERGY LEVELS

Within the atomic structure of each and every *isolated* atom there are specific energy levels associated with each shell and orbiting electron, as shown in Fig. 6. The energy levels associated with each shell will be different for every element. However, in general:

The farther an electron is from the nucleus, the higher is the energy state, and any electron that has left its parent atom has a higher energy state than any electron in the atomic structure.

Note in Fig. 6a that only specific energy levels can exist for the electrons in the atomic structure of an isolated atom. The result is a series of gaps between allowed energy levels

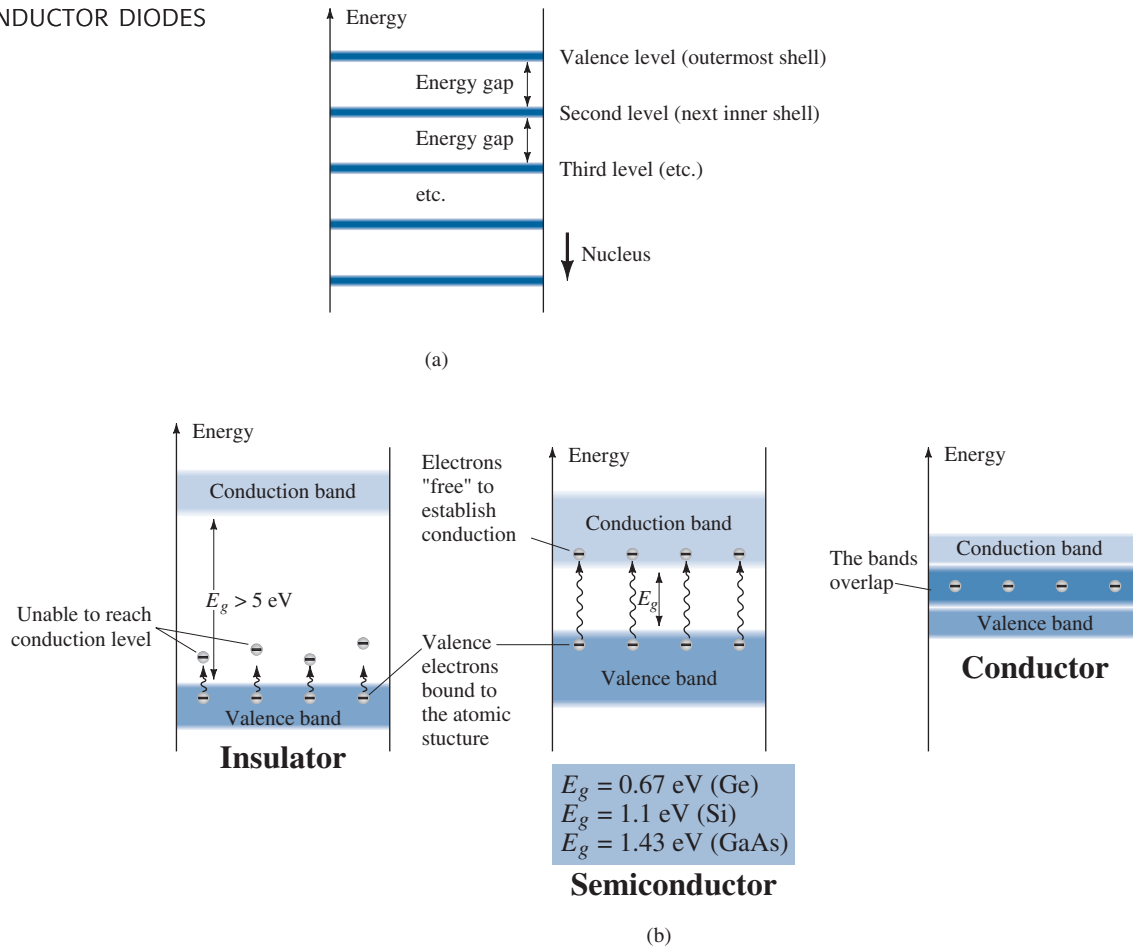


FIG. 6

Energy levels: (a) discrete levels in isolated atomic structures; (b) conduction and valence bands of an insulator, a semiconductor, and a conductor.

where carriers are not permitted. However, as the atoms of a material are brought closer together to form the crystal lattice structure, there is an interaction between atoms, which will result in the electrons of a particular shell of an atom having slightly different energy levels from electrons in the same orbit of an adjoining atom. The result is an expansion of the fixed, discrete energy levels of the valence electrons of Fig. 6a to bands as shown in Fig. 6b. In other words, the valence electrons in a silicon material can have varying energy levels as long as they fall within the band of Fig. 6b. Figure 6b clearly reveals that there is a minimum energy level associated with electrons in the conduction band and a maximum energy level of electrons bound to the valence shell of the atom. Between the two is an energy gap that the electron in the valence band must overcome to become a free carrier. That energy gap is different for Ge, Si, and GaAs; Ge has the smallest gap and GaAs the largest gap. In total, this simply means that:

An electron in the valence band of silicon must absorb more energy than one in the valence band of germanium to become a free carrier. Similarly, an electron in the valence band of gallium arsenide must gain more energy than one in silicon or germanium to enter the conduction band.

This difference in energy gap requirements reveals the sensitivity of each type of semiconductor to changes in temperature. For instance, as the temperature of a Ge sample increases, the number of electrons that can pick up thermal energy and enter the conduction band will increase quite rapidly because the energy gap is quite small. However, the number of electrons entering the conduction band for Si or GaAs would be a great deal less. This sensitivity to changes in energy level can have positive and negative effects. The design of photodetectors sensitive to light and security systems sensitive to heat would appear to be an excellent area of application for Ge devices. However, for transistor networks, where stability is a high priority, this sensitivity to temperature or light can be a detrimental factor.

The energy gap also reveals which elements are useful in the construction of light-emitting devices such as light-emitting diodes (LEDs), which will be introduced shortly. The wider the energy gap, the greater is the possibility of energy being released in the form of visible or invisible (infrared) light waves. For conductors, the overlapping of valence and conduction bands essentially results in all the additional energy picked up by the electrons being dissipated in the form of heat. Similarly, for Ge and Si, because the energy gap is so small, most of the electrons that pick up sufficient energy to leave the valence band end up in the conduction band, and the energy is dissipated in the form of heat. However, for GaAs the gap is sufficiently large to result in significant light radiation. For LEDs (Section 9) the level of doping and the materials chosen determine the resulting color.

Before we leave this subject, it is important to underscore the importance of understanding the units used for a quantity. In Fig. 6 the units of measurement are *electron volts* (eV). The unit of measure is appropriate because W (energy) = QV (as derived from the defining equation for voltage: $V = W/Q$). Substituting the charge of one electron and a potential difference of 1 V results in an energy level referred to as one *electron volt*.

That is,

$$\begin{aligned} W &= QV \\ &= (1.6 \times 10^{-19} \text{ C})(1 \text{ V}) \\ &= 1.6 \times 10^{-19} \text{ J} \end{aligned}$$

and

$$\boxed{1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}} \quad (1)$$

5 *n*-TYPE AND *p*-TYPE MATERIALS

Because Si is the material used most frequently as the base (substrate) material in the construction of solid-state electronic devices, the discussion to follow in this and the next few sections deals solely with Si semiconductors. Because Ge, Si, and GaAs share a similar covalent bonding, the discussion can easily be extended to include the use of the other materials in the manufacturing process.

As indicated earlier, the characteristics of a semiconductor material can be altered significantly by the addition of specific impurity atoms to the relatively pure semiconductor material. These impurities, although only added at 1 part in 10 million, can alter the band structure sufficiently to totally change the electrical properties of the material.

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

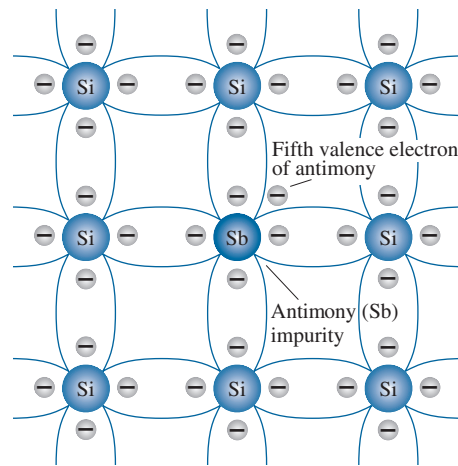
There are two extrinsic materials of immeasurable importance to semiconductor device fabrication: *n*-type and *p*-type materials. Each is described in some detail in the following subsections.

n-Type Material

Both *n*-type and *p*-type materials are formed by adding a predetermined number of impurity atoms to a silicon base. An *n*-type material is created by introducing impurity elements that have *five* valence electrons (*pentavalent*), such as *antimony*, *arsenic*, and *phosphorus*. Each is a member of a subset group of elements in the Periodic Table of Elements referred to as Group V because each has five valence electrons. The effect of such impurity elements is indicated in Fig. 7 (using antimony as the impurity in a silicon base). Note that the four covalent bonds are still present. There is, however, an additional fifth electron due to the impurity atom, which is *unassociated* with any particular covalent bond. This remaining electron, loosely bound to its parent (antimony) atom, is relatively free to move within the newly formed *n*-type material. Since the inserted impurity atom has donated a relatively “free” electron to the structure:

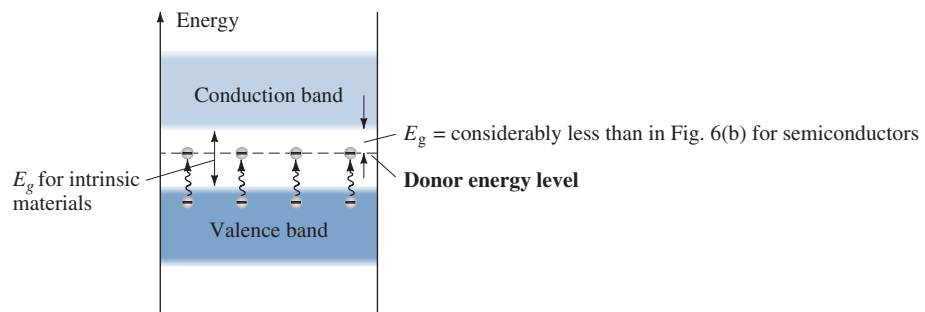
Diffused impurities with five valence electrons are called donor atoms.

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

**FIG. 7**

Antimony impurity in n-type material.

The effect of this doping process on the relative conductivity can best be described through the use of the energy-band diagram of Fig. 8. Note that a discrete energy level (called the *donor level*) appears in the forbidden band with an E_g significantly less than that of the intrinsic material. Those free electrons due to the added impurity sit at this energy level and have less difficulty absorbing a sufficient measure of thermal energy to move into the conduction band at room temperature. The result is that at room temperature, there are a large number of carriers (electrons) in the conduction level, and the conductivity of the material increases significantly. At room temperature in an intrinsic Si material there is about one free electron for every 10^{12} atoms. If the dosage level is 1 in 10 million (10^7), the ratio $10^{12}/10^7 = 10^5$ indicates that the carrier concentration has increased by a ratio of 100,000:1.

**FIG. 8**

Effect of donor impurities on the energy band structure.

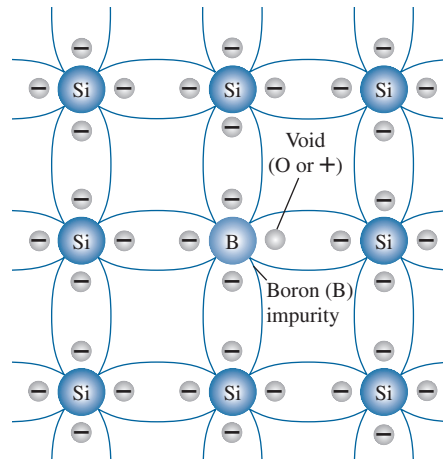
p-Type Material

The *p*-type material is formed by doping a pure germanium or silicon crystal with impurity atoms having *three* valence electrons. The elements most frequently used for this purpose are *boron*, *gallium*, and *indium*. Each is a member of a subset group of elements in the Periodic Table of Elements referred to as Group III because each has three valence electrons. The effect of one of these elements, boron, on a base of silicon is indicated in Fig. 9.

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

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

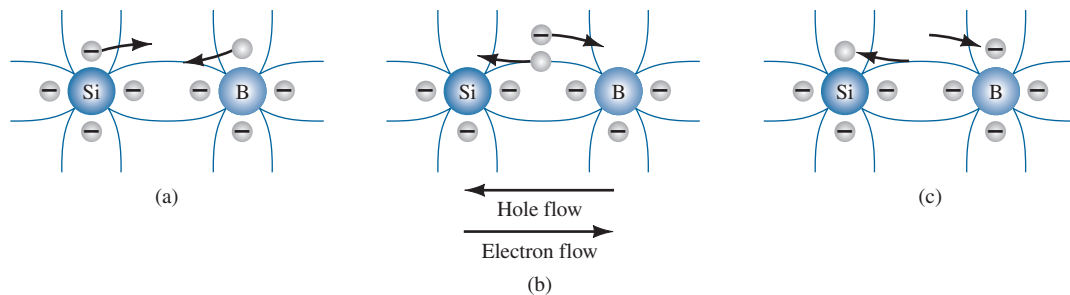
The resulting *p*-type material is electrically neutral, for the same reasons described for the *n*-type material.

**FIG. 9**

Boron impurity in p-type material.

Electron versus Hole Flow

The effect of the hole on conduction is shown in Fig. 10. If a valence electron acquires sufficient kinetic energy to break its covalent bond and fills the void created by a hole, then a vacancy, or hole, will be created in the covalent bond that released the electron. There is, therefore, a transfer of holes to the left and electrons to the right, as shown in Fig. 10. The direction to be used in this text is that of *conventional flow*, which is indicated by the direction of hole flow.

**FIG. 10**

Electron versus hole flow.

Majority and Minority Carriers

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

In an n-type material (Fig. 11a) the electron is called the majority carrier and the hole the minority carrier.

For the *p*-type material the number of holes far outweighs the number of electrons, as shown in Fig. 11b. Therefore:

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

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

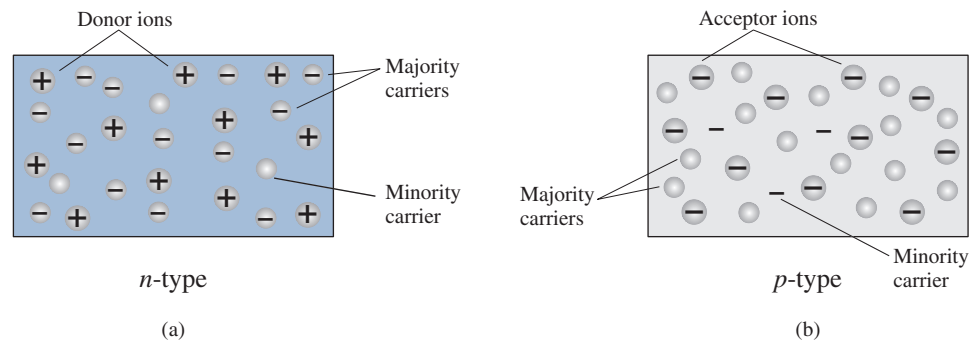


FIG. 11

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

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

6 SEMICONDUCTOR DIODE

Now that both *n*- and *p*-type materials are available, we can construct our first solid-state electronic device: The *semiconductor diode*, with applications too numerous to mention, is created by simply joining an *n*-type and a *p*-type material together, nothing more, just the joining of one material with a majority carrier of electrons to one with a majority carrier of holes. The basic simplicity of its construction simply reinforces the importance of the development of this solid-state era.

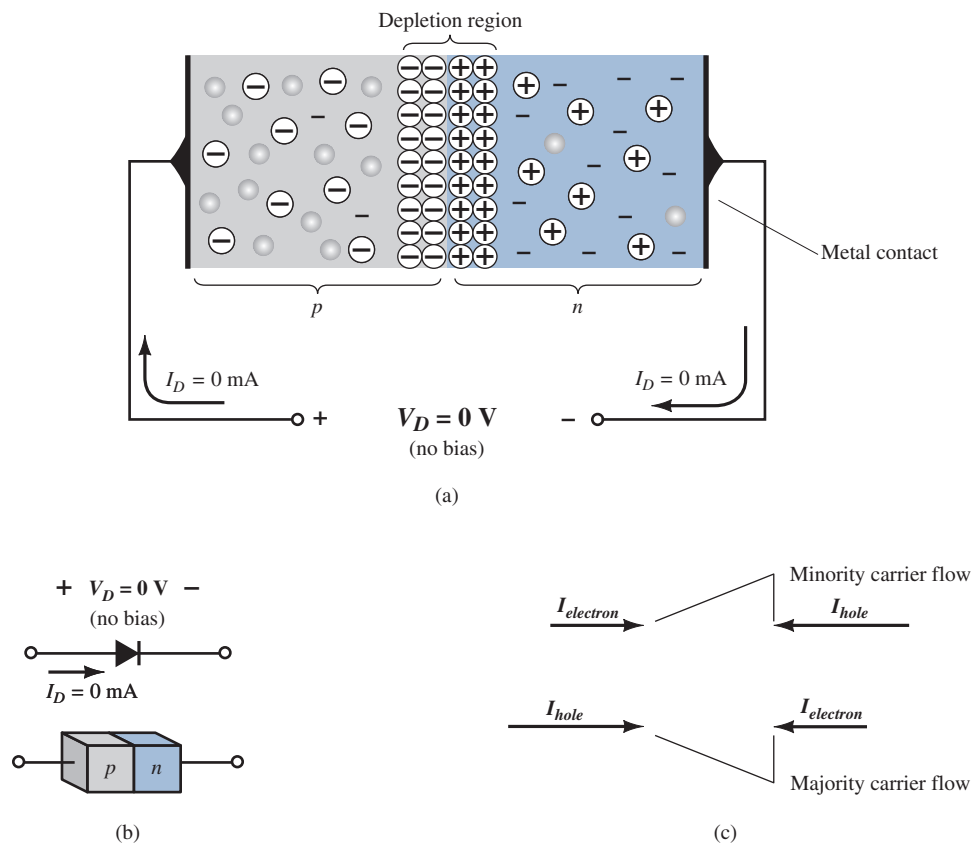
No Applied Bias ($V = 0$ V)

At the instant the two materials are “joined” the electrons and the holes in the region of the junction will combine, resulting in a lack of free carriers in the region near the junction, as shown in Fig. 12a. Note in Fig. 12a that the only particles displayed in this region are the positive and the negative ions remaining once the free carriers have been absorbed.

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

If leads are connected to the ends of each material, a *two-terminal device* results, as shown in Figs. 12a and 12b. Three options then become available: *no bias*, *forward bias*, and *reverse bias*. The term *bias* refers to the application of an external voltage across the two terminals of the device to extract a response. The condition shown in Figs. 12a and 12b is the no-bias situation because there is no external voltage applied. It is simply a diode with two leads sitting isolated on a laboratory bench. In Fig. 12b the symbol for a semiconductor diode is provided to show its correspondence with the *p*–*n* junction. In each figure it is clear that the applied voltage is 0 V (no bias) and the resulting current is 0 A, much like an isolated resistor. The absence of a voltage across a resistor results in zero current through it. Even at this early point in the discussion it is important to note the polarity of the voltage across the diode in Fig. 12b and the direction given to the current. Those polarities will be recognized as the *defined polarities* for the semiconductor diode. If a voltage applied across the diode has the same polarity across the diode as in Fig. 12b, it will be considered a positive voltage. If the reverse, it is a negative voltage. The same standards can be applied to the defined direction of current in Fig. 12b.

Under no-bias conditions, any minority carriers (holes) in the *n*-type material that find themselves within the depletion region for any reason whatsoever will pass quickly into the *p*-type material. The closer the minority carrier is to the junction, the greater is the attraction for the layer of negative ions and the less is the opposition offered by the positive ions in the depletion region of the *n*-type material. We will conclude, therefore, for future discussions, that any minority carriers of the *n*-type material that find themselves in the depletion region will pass directly into the *p*-type material. This carrier flow is indicated at the top of Fig. 12c for the minority carriers of each material.

**FIG. 12**

A p - n junction with no external bias: (a) an internal distribution of charge; (b) a diode symbol, with the defined polarity and the current direction; (c) demonstration that the net carrier flow is zero at the external terminal of the device when $V_D = 0\text{ V}$.

The majority carriers (electrons) of the n -type material must overcome the attractive forces of the layer of positive ions in the n -type material and the shield of negative ions in the p -type material to migrate into the area beyond the depletion region of the p -type material. However, the number of majority carriers is so large in the n -type material that there will invariably be a small number of majority carriers with sufficient kinetic energy to pass through the depletion region into the p -type material. Again, the same type of discussion can be applied to the majority carriers (holes) of the p -type material. The resulting flow due to the majority carriers is shown at the bottom of Fig. 12c.

A close examination of Fig. 12c will reveal that the relative magnitudes of the flow vectors are such that the net flow in either direction is zero. This cancellation of vectors for each type of carrier flow is indicated by the crossed lines. The length of the vector representing hole flow is drawn longer than that of electron flow to demonstrate that the two magnitudes need not be the same for cancellation and that the doping levels for each material may result in an unequal carrier flow of holes and electrons. In summary, therefore:

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

In other words, the current under no-bias conditions is zero, as shown in Figs. 12a and 12b.

Reverse-Bias Condition ($V_D < 0\text{ V}$)

If an external potential of V volts is applied across the p - n junction such that the positive terminal is connected to the n -type material and the negative terminal is connected to the p -type material as shown in Fig. 13, the number of uncovered positive ions in the depletion region of the n -type material will increase due to the large number of free electrons drawn to the positive potential of the applied voltage. For similar reasons, the number of uncovered negative ions will increase in the p -type material. The net effect, therefore, is a

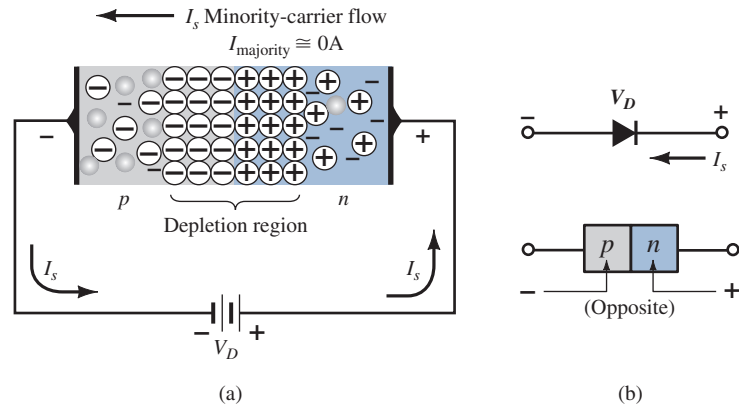


FIG. 13

Reverse-biased $p-n$ junction: (a) internal distribution of charge under reverse-bias conditions; (b) reverse-bias polarity and direction of reverse saturation current.

widening of the depletion region. This widening of the depletion region will establish too great a barrier for the majority carriers to overcome, effectively reducing the majority carrier flow to zero, as shown in Fig. 13a.

The number of minority carriers, however, entering the depletion region will not change, resulting in minority-carrier flow vectors of the same magnitude indicated in Fig. 12c with no applied voltage.

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

The reverse saturation current is seldom more than a few microamperes and typically in nA, except for high-power devices. The term *saturation* comes from the fact that it reaches its maximum level quickly and does not change significantly with increases in the reverse-bias potential, as shown on the diode characteristics of Fig. 15 for $V_D < 0$ V. The reverse-biased conditions are depicted in Fig. 13b for the diode symbol and $p-n$ junction. Note, in particular, that the direction of I_s is against the arrow of the symbol. Note also that the negative side of the applied voltage is connected to the p-type material and the positive side to the n-type material, the difference in underlined letters for each region revealing a reverse-bias condition.

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

A *forward-bias* or “on” condition is established by applying the positive potential to the p -type material and the negative potential to the n -type material as shown in Fig. 14.

The application of a forward-bias potential V_D will “pressure” electrons in the n -type material and holes in the p -type material to recombine with the ions near the boundary and reduce the width of the depletion region as shown in Fig. 14a. The resulting minority-carrier flow

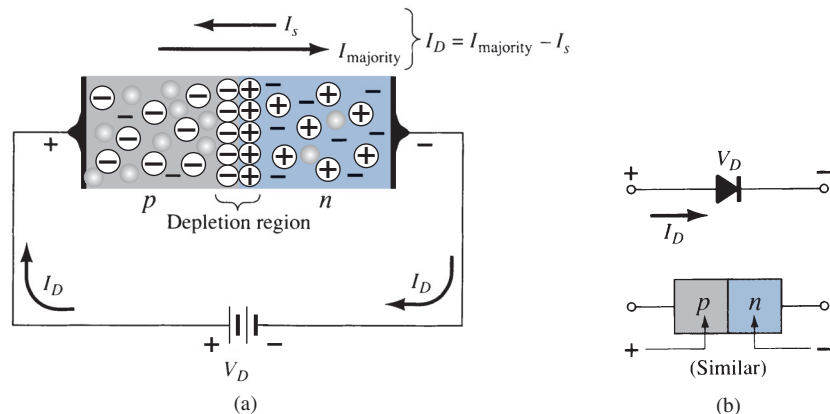


FIG. 14

Forward-biased $p-n$ junction: (a) internal distribution of charge under forward-bias conditions; (b) forward-bias polarity and direction of resulting current.

of electrons from the p -type material to the n -type material (and of holes from the n -type material to the p -type material) has not changed in magnitude (since the conduction level is controlled primarily by the limited number of impurities in the material), but the reduction in the width of the depletion region has resulted in a heavy majority flow across the junction. An electron of the n -type material now “sees” a reduced barrier at the junction due to the reduced depletion region and a strong attraction for the positive potential applied to the p -type material. As the applied bias increases in magnitude, the depletion region will continue to decrease in width until a flood of electrons can pass through the junction, resulting in an exponential rise in current as shown in the forward-bias region of the characteristics of Fig. 15. Note that the vertical scale of Fig. 15 is measured in milliamperes (although some semiconductor diodes have a vertical scale measured in amperes), and the horizontal scale in the forward-bias region has a maximum of 1 V. Typically, therefore, the voltage across a forward-biased diode will be less than 1 V. Note also how quickly the current rises beyond the knee of the curve.

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

$$I_D = I_s(e^{V_D/nV_T} - 1) \quad (A) \quad (2)$$

where I_s is the reverse saturation current
 V_D is the applied forward-bias voltage across the diode
 n is an ideality factor, which is a function of the operating conditions and physical construction; it has a range between 1 and 2 depending on a wide variety of factors ($n = 1$ will be assumed throughout this text unless otherwise noted).

The voltage V_T in Eq. (1) is called the *thermal voltage* and is determined by

$$V_T = \frac{kT_K}{q} \quad (V) \quad (3)$$

where k is Boltzmann’s constant = 1.38×10^{-23} J/K
 T_K is the absolute temperature in kelvins = $273 +$ the temperature in $^{\circ}\text{C}$
 q is the magnitude of electronic charge = 1.6×10^{-19} C

EXAMPLE 1 At a temperature of 27°C (common temperature for components in an enclosed operating system), determine the thermal voltage V_T .

Solution: Substituting into Eq. (3), we obtain

$$\begin{aligned} T &= 273 + ^{\circ}\text{C} = 273 + 27 = 300 \text{ K} \\ V_T &= \frac{kT_K}{q} = \frac{(1.38 \times 10^{-23} \text{ J/K})(30 \text{ K})}{1.6 \times 10^{-19} \text{ C}} \\ &= 25.875 \text{ mV} \cong 26 \text{ mV} \end{aligned}$$

The thermal voltage will become an important parameter in the analysis to follow in this chapter.

Initially, Eq. (2) with all its defined quantities may appear somewhat complex. However, it will not be used extensively in the analysis to follow. It is simply important at this point to understand the source of the diode characteristics and which factors affect its shape.

A plot of Eq. (2) with $I_s = 10$ pA is provided in Fig. 15 as the dashed line. If we expand Eq. (2) into the following form, the contributing component for each region of Fig. 15 can be described with increased clarity:

$$I_D = I_s e^{V_D/nV_T} - I_s$$

For positive values of V_D the first term of the above equation will grow very quickly and totally overpower the effect of the second term. The result is the following equation, which only has positive values and takes on the exponential format e^x appearing in Fig. 16:

$$I_D \cong I_s e^{V_D/nV_T} \quad (V_D \text{ positive})$$

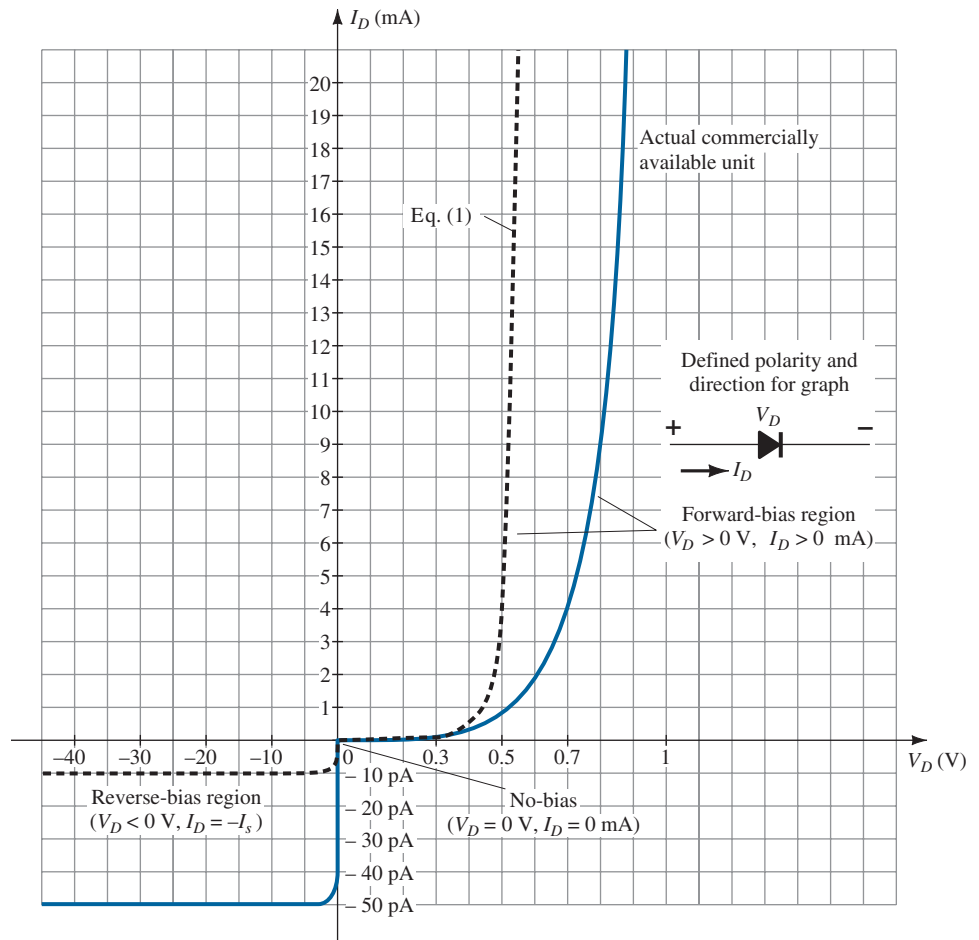


FIG. 15
Silicon semiconductor diode characteristics.

The exponential curve of Fig. 16 increases very rapidly with increasing values of x . At $x = 0$, $e^0 = 1$, whereas at $x = 5$, it jumps to greater than 148. If we continued to $x = 10$, the curve jumps to greater than 22,000. Clearly, therefore, as the value of x increases, the curve becomes almost vertical, an important conclusion to keep in mind when we examine the change in current with increasing values of applied voltage.

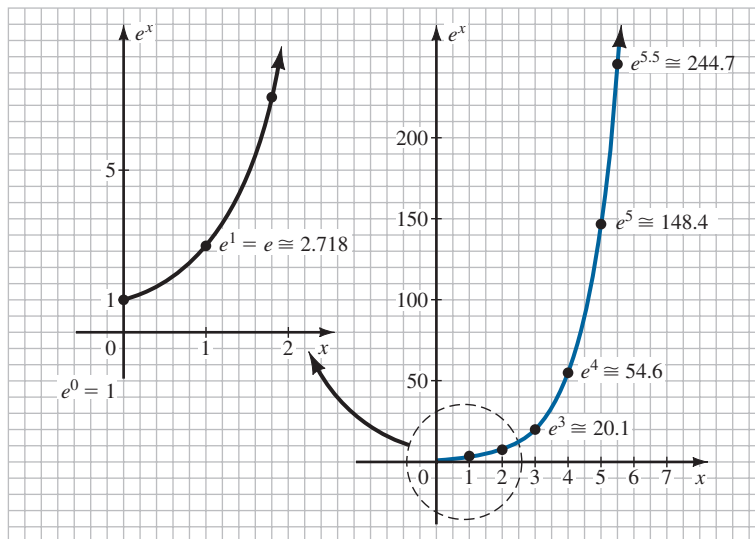


FIG. 16
Plot of e^x .

For negative values of V_D the exponential term drops very quickly below the level of I_s and the resulting equation for I_D is simply

$$I_D \cong -I_s \quad (V_D \text{ negative})$$

Note in Fig. 15 that for negative values of V_D the current is essentially horizontal at the level of $-I_s$.

At $V = 0$ V, Eq. (2) becomes

$$I_D = I_s(e^0 - 1) = I_s(1 - 1) = 0 \text{ mA}$$

as confirmed by Fig. 15.

The sharp change in direction of the curve at $V_D = 0$ V is simply due to the change in current scales from above the axis to below the axis. Note that above the axis the scale is in milliamperes (mA), whereas below the axis it is in picoamperes (pA).

Theoretically, with all things perfect, the characteristics of a silicon diode should appear as shown by the dashed line of Fig. 15. However, commercially available silicon diodes deviate from the ideal for a variety of reasons including the internal “body” resistance and the external “contact” resistance of a diode. Each contributes to an additional voltage at the same current level, as determined by Ohm’s law, causing the shift to the right witnessed in Fig. 15.

The change in current scales between the upper and lower regions of the graph was noted earlier. For the voltage V_D there is also a measurable change in scale between the right-hand region of the graph and the left-hand region. For positive values of V_D the scale is in tenths of volts, and for the negative region it is in tens of volts.

It is important to note in Fig. 14b how:

The defined direction of conventional current for the positive voltage region matches the arrowhead in the diode symbol.

This will always be the case for a forward-biased diode. It may also help to note that the forward-bias condition is established when the bar representing the negative side of the applied voltage matches the side of the symbol with the vertical bar.

Going back a step further by looking at Fig. 14b, we find a forward-bias condition is established across a p - n junction when the positive side of the applied voltage is applied to the p -type material (noting the correspondence in the letter p) and the negative side of the applied voltage is applied to the n -type material (noting the same correspondence).

It is particularly interesting to note that the reverse saturation current of the commercial unit is significantly larger than that of I_s in Shockley’s equation. In fact,

The actual reverse saturation current of a commercially available diode will normally be measurably larger than that appearing as the reverse saturation current in Shockley’s equation.

This increase in level is due to a wide range of factors that include

- **leakage currents**
- **generation of carriers in the depletion region**
- **higher doping levels** that result in increased levels of reverse current
- **sensitivity to the intrinsic level of carriers** in the component materials by a squared factor—double the intrinsic level, and the contribution to the reverse current could increase by a factor of four.
- **a direct relationship with the junction area**—double the area of the junction, and the contribution to the reverse current could double. High-power devices that have larger junction areas typically have much higher levels of reverse current.
- **temperature sensitivity**—for every 5°C increase in current, the level of reverse saturation current in Eq. 2 will double, whereas a 10°C increase in current will result in doubling of the actual reverse current of a diode.

Note in the above the use of the terms reverse saturation current and reverse current. The former is simply due to the physics of the situation, whereas the latter includes all the other possible effects that can increase the level of current.

We will find in the discussions to follow that the ideal situation is for I_s to be 0 A in the reverse-bias region. The fact that it is typically in the range of 0.01 pA to 10 pA today as compared to 0.1 μA to 1 μA a few decades ago is a credit to the manufacturing industry. Comparing the common value of 1 nA to the 1- μA level of years past shows an improvement factor of 100,000.

Breakdown Region

Even though the scale of Fig. 15 is in tens of volts in the negative region, there is a point where the application of too negative a voltage with the reverse polarity will result in a sharp change in the characteristics, as shown in Fig. 17. The current increases at a very rapid rate in a direction opposite to that of the positive voltage region. The reverse-bias potential that results in this dramatic change in characteristics is called the *breakdown potential* and is given the label V_{BV} .

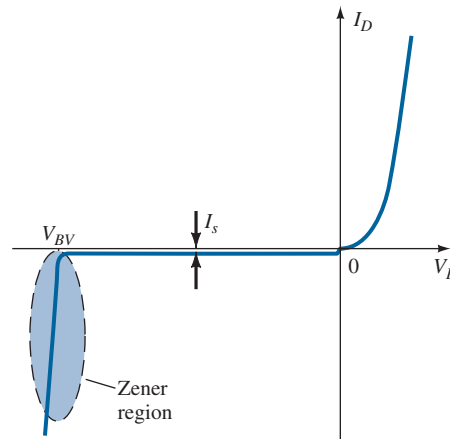


FIG. 17

Breakdown region.

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

The avalanche region (V_{BV}) can be brought closer to the vertical axis by increasing the doping levels in the *p*- and *n*-type materials. However, as V_{BV} decreases to very low levels, such as -5 V, another mechanism, called *Zener breakdown*, will contribute to the sharp change in the characteristic. It occurs because there is a strong electric field in the region of the junction that can disrupt the bonding forces within the atom and “generate” carriers. Although the Zener breakdown mechanism is a significant contributor only at lower levels of V_{BV} , this sharp change in the characteristic at any level is called the *Zener region*, and diodes employing this unique portion of the characteristic of a *p-n* junction are called *Zener diodes*. They are described in detail in Section 15.

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

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

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

In general, the breakdown voltage of GaAs diodes is about 10% higher those for silicon diodes but after 200% higher than levels for Ge diodes.

Ge, Si, and GaAs

The discussion thus far has solely used Si as the base semiconductor material. It is now important to compare it to the other two materials of importance: GaAs and Ge. A plot comparing the characteristics of Si, GaAs, and Ge diodes is provided in Fig. 18. The curves are not

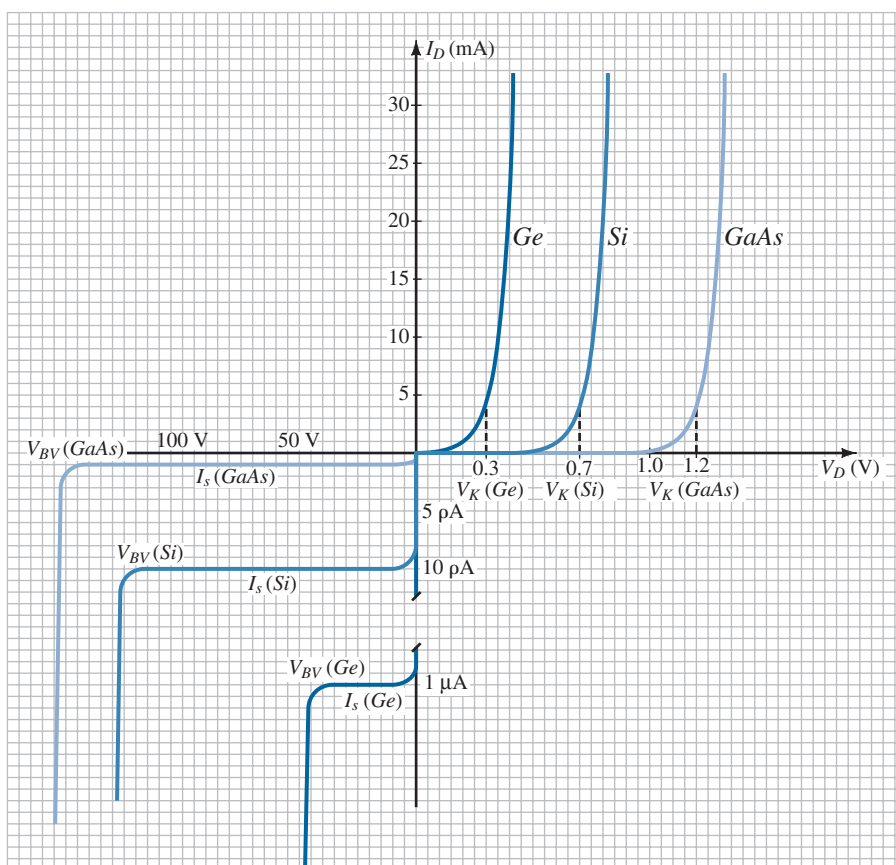


FIG. 18

Comparison of Ge, Si, and GaAs commercial diodes.

simply plots of Eq. 2 but the actual response of commercially available units. The total reverse current is shown and not simply the reverse saturation current. It is immediately obvious that the point of vertical rise in the characteristics is different for each material, although the general shape of each characteristic is quite similar. Germanium is closest to the vertical axis and GaAs is the most distant. As noted on the curves, the center of the knee (hence the K is the notation V_K) of the curve is about 0.3 V for Ge, 0.7 V for Si, and 1.2 V for GaAs (see Table 3).

The shape of the curve in the reverse-bias region is also quite similar for each material, but notice the measurable difference in the magnitudes of the typical reverse saturation currents. For GaAs, the reverse saturation current is typically about 1 pA, compared to 10 pA for Si and 1 μ A for Ge, a significant difference in levels.

Also note the relative magnitudes of the reverse breakdown voltages for each material. GaAs typically has maximum breakdown levels that exceed those of Si devices of the same power level by about 10%, with both having breakdown voltages that typically extend between 50 V and 1 kV. There are Si power diodes with breakdown voltages as high as 20 kV. Germanium typically has breakdown voltages of less than 100 V, with maximums around 400 V. The curves of Fig. 18 are simply designed to reflect relative breakdown voltages for the three materials. When one considers the levels of reverse saturation currents and breakdown voltages, Ge certainly sticks out as having the least desirable characteristics.

A factor not appearing in Fig. 18 is the operating speed for each material—an important factor in today's market. For each material, the electron mobility factor is provided in Table 4. It provides an indication of how fast the carriers can progress through the material and therefore the operating speed of any device made using the materials. Quite obviously, GaAs stands out, with a mobility factor more than five times that of silicon and twice that of germanium. The result is that GaAs and Ge are often used in high-speed applications. However, through proper design, careful control of doping levels, and so on, silicon is also found in systems operating in the gigahertz range. Research today is also looking at compounds in groups III–V that have even higher mobility factors to ensure that industry can meet the demands of future high-speed requirements.

TABLE 3
Knee Voltages V_K

Semiconductor	V_K (V)
Ge	0.3
Si	0.7
GaAs	1.2

TABLE 4
Electron Mobility μ_n

Semiconductor	μ_n ($\text{cm}^2/\text{V}\cdot\text{s}$)
Ge	3900
Si	1500
GaAs	8500

EXAMPLE 2 Using the curves of Fig 18:

- Determine the voltage across each diode at a current of 1 mA.
- Repeat for a current of 4 mA.
- Repeat for a current of 30 mA.
- Determine the average value of the diode voltage for the range of currents listed above.
- How do the average values compare to the knee voltages listed in Table 3?

Solution:

- $V_D(\text{Ge}) = 0.2 \text{ V}$, $V_D(\text{Si}) = 0.6 \text{ V}$, $V_D(\text{GaAs}) = 1.1 \text{ V}$
- $V_D(\text{Ge}) = 0.3 \text{ V}$, $V_D(\text{Si}) = 0.7 \text{ V}$, $V_D(\text{GaAs}) = 1.2 \text{ V}$
- $V_D(\text{Ge}) = 0.42 \text{ V}$, $V_D(\text{Si}) = 0.82 \text{ V}$, $V_D(\text{GaAs}) = 1.33 \text{ V}$
- Ge: $V_{\text{av}} = (0.2 \text{ V} + 0.3 \text{ V} + 0.42 \text{ V})/3 = 0.307 \text{ V}$
 Si: $V_{\text{av}} = (0.6 \text{ V} + 0.7 \text{ V} + 0.82 \text{ V})/3 = 0.707 \text{ V}$
 GaAs: $V_{\text{av}} = (1.1 \text{ V} + 1.2 \text{ V} + 1.33 \text{ V})/3 = 1.21 \text{ V}$
- Very close correspondence. Ge: 0.307 V vs. 0.3, V, Si: 0.707 V vs. 0.7 V, GaAs: 1.21 V vs. 1.2 V.

Temperature Effects

Temperature can have a marked effect on the characteristics of a semiconductor diode, as demonstrated by the characteristics of a silicon diode shown in Fig. 19:

In the forward-bias region the characteristics of a silicon diode shift to the left at a rate of 2.5 mV per centigrade degree increase in temperature.

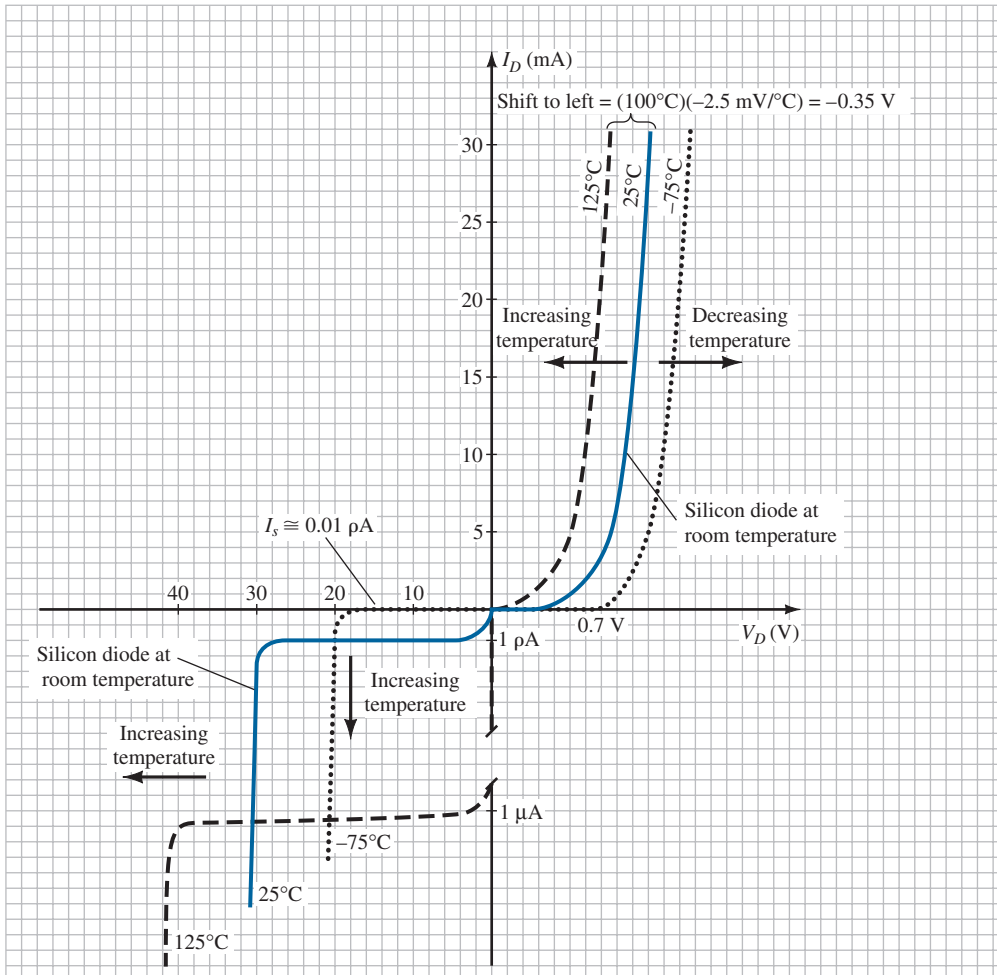


FIG. 19

Variation in Si diode characteristics with temperature change.

An increase from room temperature (20°C) to 100°C (the boiling point of water) results in a drop of $80(2.5 \text{ mV}) = 200 \text{ mV}$, or 0.2 V, which is significant on a graph scaled in tenths of volts. A decrease in temperature has the reverse effect, as also shown in the figure:

In the reverse-bias region the reverse current of a silicon diode doubles for every 10°C rise in temperature.

For a change from 20°C to 100°C, the level of I_s increases from 10 nA to a value of 2.56 μA , which is a significant, 256-fold increase. Continuing to 200°C would result in a monstrous reverse saturation current of 2.62 mA. For high-temperature applications one would therefore look for Si diodes with room-temperature I_s closer to 10 pA, a level commonly available today, which would limit the current to 2.62 μA . It is indeed fortunate that both Si and GaAs have relatively small reverse saturation currents at room temperature. GaAs devices are available that work very well in the -200°C to $+200^\circ\text{C}$ temperature range, with some having maximum temperatures approaching 400°C. Consider, for a moment, how huge the reverse saturation current would be if we started with a Ge diode with a saturation current of 1 μA and applied the same doubling factor.

Finally, it is important to note from Fig. 19 that:

The reverse breakdown voltage of a semiconductor diode will increase or decrease with temperature.

However, if the initial breakdown voltage is less than 5 V, the breakdown voltage may actually decrease with temperature. The sensitivity of the breakdown potential to changes of temperature will be examined in more detail in Section 15.

Summary

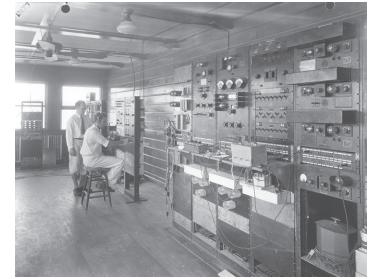
A great deal has been introduced in the foregoing paragraphs about the construction of a semiconductor diode and the materials employed. The characteristics have now been presented and the important differences between the response of the materials discussed. It is now time to compare the p - n junction response to the desired response and reveal the primary functions of a semiconductor diode.

Table 5 provides a synopsis of material regarding the three most frequently used semiconductor materials. Figure 20 includes a short biography of the first research scientist to discover the p - n junction in a semiconductor material.

TABLE 5

The Current Commercial Use of Ge, Si, and GaAs

Ge:	Germanium is in limited production due to its temperature sensitivity and high reverse saturation current. It is still commercially available but is limited to some high-speed applications (due to a relatively high mobility factor) and applications that use its sensitivity to light and heat such as photodetectors and security systems.
Si:	Without question the semiconductor used most frequently for the full range of electronic devices. It has the advantage of being readily available at low cost and has relatively low reverse saturation currents, good temperature characteristics, and excellent breakdown voltage levels. It also benefits from decades of enormous attention to the design of large-scale integrated circuits and processing technology.
GaAs:	Since the early 1990s the interest in GaAs has grown in leaps and bounds, and it will eventually take a good share of the development from silicon devices, especially in very large scale integrated circuits. Its high-speed characteristics are in more demand every day, with the added features of low reverse saturation currents, excellent temperature sensitivities, and high breakdown voltages. More than 80% of its applications are in optoelectronics with the development of light-emitting diodes, solar cells, and other photodetector devices, but that will probably change dramatically as its manufacturing costs drop and its use in integrated circuit design continues to grow; perhaps the semiconductor material of the future.



Russell Ohl (1898–1987)

American (Allentown, PA; Holmdel, NJ; Vista, CA) Army Signal Corps, University of Colorado, Westinghouse, AT&T, Bell Labs Fellow, Institute of Radio Engineers—1955 (Courtesy of AT&T Archives History Center.)

Although vacuum tubes were used in all forms of communication in the 1930s, Russell Ohl was determined to demonstrate that the future of the field was defined by semiconductor crystals. Germanium was not immediately available for his research, so he turned to silicon, and found a way to raise its level of purity to 99.8%, for which he received a patent. The actual discovery of the p - n junction, as often happens in scientific research, was the result of a set of circumstances that were not planned. On February 23, 1940, Ohl found that a silicon crystal with a crack down the middle would produce a significant rise in current when placed near a source of light. This discovery led to further research, which revealed that the purity levels on each side of the crack were different and that a barrier was formed at the junction that allowed the passage of current in only one direction—the first solid-state diode had been identified and explained. In addition, this sensitivity to light was the beginning of the development of solar cells. The results were quite instrumental in the development of the transistor in 1945 by three individuals also working at Bell Labs.

FIG. 20

7 IDEAL VERSUS PRACTICAL

In the previous section we found that a p - n junction will permit a generous flow of charge when forward-biased and a very small level of current when reverse-biased. Both conditions are reviewed in Fig. 21, with the heavy current vector in Fig. 21a matching the direction of the arrow in the diode symbol and the significantly smaller vector in the opposite direction in Fig. 21b representing the reverse saturation current.

An analogy often used to describe the behavior of a semiconductor diode is a mechanical switch. In Fig. 21a the diode is acting like a closed switch permitting a generous flow of charge in the direction indicated. In Fig. 21b the level of current is so small in most cases that it can be approximated as 0 A and represented by an open switch.

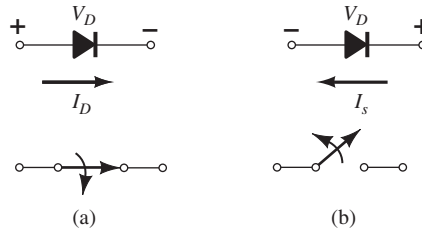


FIG. 21

Ideal semiconductor diode: (a) forward-biased; (b) reverse-biased.

In other words:

The semiconductor diode behaves in a manner similar to a mechanical switch in that it can control whether current will flow between its two terminals.

However, it is important to also be aware that:

The semiconductor diode is different from a mechanical switch in the sense that when the switch is closed it will only permit current to flow in one direction.

Ideally, if the semiconductor diode is to behave like a closed switch in the forward-bias region, the resistance of the diode should be $0\ \Omega$. In the reverse-bias region its resistance should be $\infty\ \Omega$ to represent the open-circuit equivalent. Such levels of resistance in the forward- and reverse-bias regions result in the characteristics of Fig. 22.

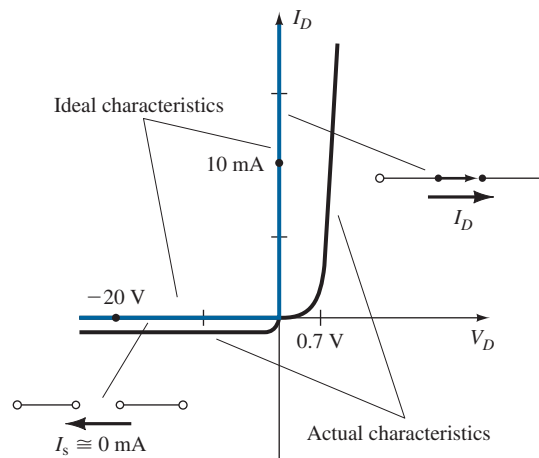


FIG. 22

Ideal versus actual semiconductor characteristics.