DOCUMENT RESUME

ED 275 864 CE 045 460

AUTHOR	Sutliff, Ronald D.; And Others
TITLE	Solid-State Devices.
INSTITUTION	Marine Corps Inst., Washington, DC.
REPORT NO	MCI-11-42
PUB DATE	1 Aug 86
NOTE	173p.; Document is set in small type. Some graphics may not reproduce clearly.
PUB TYPE	Guides - Classroom Use - Materials (For Learner) (051)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Autoinstructional Aids; Continuing Education; Course Content; Electronic Equipment; *Electronics; Electronic Technicians; Extension Education; Independent Study; Learning Activities; Military Personnel; *Military Training; Postsecondary Education; *Semiconductor Devices; Skilled Occupations; *Transistors
IDENTIFIERS	Marine Corps; Military Curriculum Materials; *Solid State (Electronics)

ABSTRACT

This self-study course is designed to familiarize Marine Corps enlisted personnel with the principles of solid-state devices and their functions. The course contains four study units. Each study unit begins with a general objective, which is a statement of what the student should learn from the unit. The study units are divided into numbered work units, each presenting one or more specific objectives. Text is furnished, illustrated as needed, for each work unit. At the end of the work units are study questions, with answers listed at the end of the study unit. A review lesson completes the course. The four units of the course cover the following subjects: semiconductor diodes; transistors; special devices; and solid-state power supplies. (KC)

UNITED STATES MARINE CORPS
MARINE CORPS INSTITUTE
ARLINGTON, VA 22222-0001

IN REPLY TO
11.42
1 August 86

1. ORIGIN

MCI course 11.42, Solid-State Devices, has been prepared by the Marine Corps Institute.

2. APPLICABILITY

This course is for instructional purposes only.

R. A. MALONEY
Major, U. S. Marine Corps
Deputy Director

ACKNOWLEDGMENT

The Marine Corps Institute, Marine Barracks, Washington, D.C. gratefully acknowledges the important contributions provided by the following MCI personnel in developing and publishing this course:

Course Developer	GySgt Ronald D. Sutliff
Officer in Charge	GySgt James R. Younce
Education Specialist	Mr. Olan Martin
Course Editor	Ms. Monica Noell
Wordprocessing Technician(s)	MSgt Donald L. Hamilton
. . . . .	GySgt Ronald D. Sutliff
. . . . .	Cpl S. R. Armstead
. . . . .	LCpl M. Howe
. . . . .	LCpl M. A. Gallo

The Marine Corps Institute gratefully acknowledges the assistance of Engineer Maintenance Company, 2nd Maintenance Battalion, 2nd Force Service Support Group, Camp Lejeune, North Carolina, in the validation of this course.

The Marine Corps Institute gratefully acknowledges the assistance of Marine Corps Engineer Schools, Marine Corps Base Camp Lejeune, North Carolina, in the formal review of this course.

INFORMATION

FOR

MCI STUDENTS

Welcome to the Marine Corps Institute training program. Your interest in self-improvement and increased professional competence is commendable.

Information is provided below to assist you in completing the course. Please read this guidance before proceeding with your studies.

1. MATERIALS

Check your course materials. You should have all the materials listed in the "Course Introduction." In addition you should have an envelope to mail your review lesson back to MCI for grading unless your review lesson answer sheet is of the self-mailing type. If your answer sheet is the pre-printed type, check to see that your name, rank, and social security number are correct. Check closely, your MCI records are kept on a computer and any discrepancy in the above information may cause your subsequent activity to go unrecorded. You may correct the information directly on the answer sheet. If you did not receive all your materials, notify your training NCO. If you are not attached to a Marine Corps unit, request them through the Hotline (autovon 288-4175 or commercial 202-433-4175).

2. LESSON SUBMISSION

The self-graded exercises contained in your course are not to be returned to MCI. Only the completed review lesson answer sheet should be mailed to MCI. The answer sheet is to be completed and mailed only after you have finished all of the study units in the course booklet. The review lesson has been designed to prepare you for the final examination.

It is important that you provide the required information at the bottom of your review lesson answer sheet if it does not have your name and address printed on it. In courses in which the work is submitted on blank paper or printed forms, identify each sheet in the following manner:

DOE, John J. Sgt 332-11-9999
08.4g, Forward Observation
Review Lesson
Military or office address
(RUC number, if available)

Submit your review lesson on the answer sheet and/or forms provided. Complete all blocks and follow the directions on the answer sheet for mailing. Otherwise, your answer sheet may be delayed or lost. If you have to interrupt your studies for any reason and find that you cannot complete your course in one year, you may request a single six month extension by contacting your training NCO, at least one month prior to your course completion deadline date. If you are not attached to a Marine Corps unit you may make this request by letter. Your commanding officer is notified monthly of your status through the monthly Unit Activity Report. In the event of difficulty, contact your training NCO or MCI immediately.

3. MAIL-TIME DELAY

Presented below are the mail-time delays that you may experience between the mailing of your review lesson and its return to you.

	TURNAROUND MAIL TIME	MCI PROCESSING TIME	TOTAL NUMBER DAYS
EAST COAST	16	5	21
WEST COAST	16	5	21
FPO NEW YORK	18	5	23
FPO SAN FRANCISCO	22	5	27

You may also experience a short delay in receiving your final examination due to administrative screening required at MCI.

4. GRADING SYSTEM

LESSONS			EXAMS	
GRADE	PERCENT	MEANING	GRADE	PERCENT
A	94-100	EXCELLENT	A	94-100
B	86-93	ABOVE AVERAGE	B	86-93
C	78-85	AVERAGE	C	78-85
D	70-77	BELOW AVERAGE	D	65-77
NL	BELOW 70	FAILING	F	BELOW 65

You will receive a percentage grade for your review lesson and for the final examination. A review lesson which receives a score below 70 is given a grade of NL (no lesson). It must be resubmitted and PASSED before you will receive an examination. The grade attained on the final exam is your course grade, unless you fail your first exam. Those who fail their first exam will be sent an alternate exam in which the highest grade possible is 65%. Failure of the alternate will result in failure of the course.

5. FINAL EXAMINATION

ACTIVE DUTY PERSONNEL: When you pass your REVIEW LESSON, your examination will be mailed automatically to your commanding officer. The administration of MCI final examinations must be supervised by a commissioned or warrant officer or a staff NCO.

OTHER PERSONNEL: Your examination may be administered and supervised by your supervisor.

6. COMPLETION CERTIFICATE

The completion certificate will be mailed to your commanding officer and your official records will be updated automatically. For non Marines, your completion certificate is mailed to your supervisor.

7. RESERVE RETIREMENT CREDITS

Reserve retirement credits are awarded to inactive duty personnel only. Credits awarded for each course are listed in the "Course Introduction." Credits are only awarded upon successful completion of the course. Reserve retirement credits are not awarded for MCI study performed during drill periods if credits are also awarded for drill attendance.

8. DISENROLLMENT

Only your commanding officer can request your disenrollment from an MCI course. However, an automatic disenrollment occurs if the course is not completed (including the final exam) by the time you reach the CCD (course completion deadline) or the ACCD (adjusted course completion deadline) date. This action will adversely affect the unit's completion rate.

9. ASSISTANCE

Consult your training NCO if you have questions concerning course content. Should he/she be unable to assist you, MCI is ready to help you whenever you need it. Please use the Student Course Content Assistance Request Form (ISD-1) attached to the end of your course booklet or call one of the AUTOVON telephone numbers listed below for the appropriate course writer section.

PERSONNEL/ADMINISTRATION	288-3259
COMMUNICATIONS/ELECTRONICS/AVIATION	
NBC/INTELLIGENCE	288-3604
INFANTRY	288-3611
ENGINEER/MOTOR TRANSPORT	288-2275
SUPPLY/FOOD SERVICES/FISCAL	288-2285
TANKS/ARTILLERY/INFANTPY WEAPONS REPAIR	
LOGISTICS/EMBARKATION/MAINTENANCE MANAGEMENT/	
ASSAULT AMPHIBIAN VEHICLES	288-2290

For administrative problems use the UAR or call the MCI HOTLINE: 288-4175.

For commercial phone lines, use area code 202 and prefix 433 instead of 288.

SOLID-STATE DEVICES

Course Introduction

SOLID-STATE DEVICES is designed to familiarize the student with the principles of solid-state devices and their functions.

ADMINISTRATIVE INFORMATION

ORDER OF STUDIES

Study Unit Number	Study Hours	Subject Matter
1	3	Semiconductor Diodes
2	3	Transistors
3	3	Special Devices
4	4	Solid-State Power Supplies
	2	REVIEW LESSON
	2	FINAL EXAMINATION
	17	

- -

RESERVE RETIREMENT CREDITS:	6
EXAMINATION:	Supervised final examination without text or notes with a time limit of 2 hours
MATERIALS:	MCI 11.42, Solid-State Devices, review lesson and answer sheet.
RETURN OF MATERIALS:	Students who successfully complete this course are permitted to keep the course materials. Students disenrolled for inactivity or at the request of their commanding officer will return all course materials.

- -

SOURCE MATERIALS

FM 11-62	Solid-State Devices and Solid-State Power Supplies, Sept 83
ELECTRICITY AND ELECTRONICS	1980, 5th edition

- -

HOW TO TAKE THIS COURSE

This course contains 4 study units. Each study unit begins with a general objective that is a statement of what you should learn from the study unit. The study units are divided into numbered work units, each presenting one or more specific objectives. Read the objective(s) and then the work unit text. At the end of the work unit are study questions that you should be able to answer without referring to the text of the work unit. After answering the questions, check your answers against the correct ones listed at the end of the study unit. If you miss any of the questions, you should restudy the text of the work unit until you understand the correct responses. When you have mastered one study unit, move on to the next. After you have completed all study units, complete the review lesson and take it to your training officer or NCO for mailing to MCI. MCI will mail the final examination to your training officer or NCO when you pass the review lesson.

TABLE OF CONTENTS

Study Unit 1. SEMICONDUCTOR DIODES

Study Unit 2. TRANSISTORS

Study Unit 3. SPECIAL DEVICES

Study Unit 4. SOLID-STATE POWER SUPPLIES

MARINE CORPS INSTITUTE STUDY GUIDE

Congratulations for enrolling in the Marine Corps Institute's correspondence training program! By enrolling in this program, you have shown a desire to improve the skills you need to enhance your on-the-job performance.

Since 1920, MCI has been helping tens of thousands of hard-charging young Marines, like yourself, achieve educational goals by teaching necessary new skills or reinforcing existing skills. MCI will do every thing possible to help you reach your individual goals, whatever they may be.

Before you begin your course of instruction, you may be asking yourself, "How much will I benefit from a correspondence training program?" The answer to this depends upon you, *"YOUR PROFESSIONAL TRAITS"* (what you bring to the learning situation).

Because you have enrolled in an MCI course, your professional traits are evident and we know that:

YOU ARE PROPERLY MOTIVATED. You made a positive decision to get training on your own. Self-motivation is perhaps the most important force in learning-or achieving—anything. Wanting to learn something badly enough so that you will do what's necessary to learn—*THAT IS MOTIVATION.*

YOU SEEK TO IMPROVE YOURSELF. You enrolled to learn new skills and develop special abilities.

YOU HAVE THE INITIATIVE TO ACT. By acting on your own, you have shown that you are a self-starter, willing to reach out for opportunities.

YOU ACCEPT CHALLENGES. You have self-confidence and believe in your ability to gain training in your areas of interest.

YOU ARE ABLE TO SET PRACTICAL GOALS. You are willing to commit time, effort, and resources toward accomplishing what you set out to do. These professional traits will help you achieve success in your MCI program.

To begin your course of study:

* Look at the course introduction page. Read the COURSE INTRODUCTION to get the "nitty gritty" of what the course is about. Then read the MATERIALS section near the bottom of the page to find out which text(s) and study aids you should have received with the course. If any of the listed materials are missing, see *Information for MCI Students* to find out how to obtain them. If you have everything that is listed, you are ready to begin your MCI course.

* Read through the *TABLE OF CONTENTS* of your text(s). Note the various subjects covered in the course and the order in which they are taught. Leaf through the text(s) and look at the illustrations. Read a few work unit exercise questions to get an idea of the types of questions that are asked. If MCI provides other study aids, such as a slide rule or a plotting board, familiarize yourself with them. Now, you are ready to begin work on your MCI course.

* Turn to the first page of study unit 1. On this page you will find the study unit goal. This is a statement of what you should be able to do when you complete the final exam. Each study unit is divided into work units. Each work unit contains one terminal learning objective and several enabling objectives. The terminal learning objective is what you should be able to accomplish when you complete the work unit exercises. The enabling objectives are the steps you need to learn to help you accomplish the terminal learning objective. Read each objective for the work unit and then read the work unit text carefully. Make notes on the ideas you feel are important.

* Without referring to the text, answer the questions in each exercise.

* Check your answers against the correct ones listed at the end of the study unit.

* If you miss any of the questions, restudy the work unit until you understand the correct response.

* Go on to the next work unit, repeating the above steps, until you have completed all the work units in the study unit.

* Follow the same procedure for each study unit of the course. If you have problems with the text or work unit questions that you cannot solve on your own, ask your training NCO for the name of someone who can help you. If they cannot aid you, request assistance from MCI on the Student Course Content Assistance Request included with this course, or refer to your INFORMATION FOR MCI STUDENTS (MCI-R24i-NRL) for the telephone number of the appropriate Course Developing Division at MCI.

* When you have finished all the study units, complete the course review lesson. Try to answer each question without the aid of reference materials. However, if you do not know an answer, look it up. When you have finished the review lesson, take it to your training officer or NCO for mailing to MCI. MCI will grade it and send you a feedback sheet (MCI-R69) with your final examination listing course references for any questions that you missed on the review lesson.

"RECON" Reviews:

To prepare for your final examination you *must* review what you learned in the course. Therefore, why not make reviewing as interesting as possible. The following suggestions will make reviewing not only interesting but also a challenge.

1. Challenge yourself. Reconstruct the learning event *in your mind.* Try to recall and recapture an entire learning sequence, without notes or other references. Can you do it? You just have to "look back" to see if you've left anything out, and *that* will be an interesting read-through (review) for you.

Undoubtedly, you'll find that you were not able to recall everything. But with a little effort you'll be able to recall a great deal of the information.

Also, knowing that you are going to conduct a "reconstruct-review" will change the way you approach your learning session. You will try to learn so that you will be able to "reconstruct the event."

2. Use unused minutes. While waiting at sick bay, riding in a truck or bus, living through field duty, or just waiting to muster—use these minutes to review. Read your notes or a portion of a study unit, recalculate problems, do self-checks a second time; you can do many of these things during "unused" minutes. Just thinking about a sequence of instruction will refresh your memory to help "secure" your learning.

3. Apply what you've learned. Always, it is best to do the thing you've learned. Even if you cannot immediately put the lesson to work, sometimes you can "simulate" the learning situation. For example, make up and solve your own problems. Make up problems that take you through most of the elements of a study unit.

4. Use the "shakedown cruise" technique. Ask a fellow Marine to lend a hand and have him ask you questions about the course. Give him a particular study unit and let him fire away. It can be interesting and challenging.

The point is, reviews are necessary for good learning, but they don't have to be long and tedious. Several short reviews can be very beneficial.

Semper Fi

STUDY UNIT 1

SEMICONDUCTOR DIODES

STUDY UNIT OBJECTIVE: WITHOUT THE AID OF REFERENCES, YOU WILL IDENTIFY THE DIFFERENCES BETWEEN A CONDUCTOR, AN INSULATOR, AND A SEMICONDUCTOR. YOU WILL IDENTIFY THE ELECTRON AND HOLD FLOW THEORY IN SEMICONDUCTORS AND HOW THE SEMICONDUCTOR IS AFFECTED BY DOPING. YOU WILL IDENTIFY THE "DIODE" AND HOW IT IS CONSTRUCTED AND HOW IT OPERATES. YOU WILL ALSO IDENTIFY HOW THE DIODE CAN BE USED AS A HALF-WAVE RECTIFIER AND AS A SWITCH. IN ADDITION, YOU WILL IDENTIFY THE DIODE BY ITS SYMBOLOGY, ALPHANUMERICAL DESIGNATION, AND COLOR CODE. LASTLY, YOU WILL IDENTIFY THE PRECAUTIONS THAT MUST BE TAKEN WHEN WORKING WITH DIODES AND THE DIFFERENT WAYS TO TEST THEM.

Work Unit 1-1. INTRODUCTION, DEVELOPMENT, AND APPLICATION

DEFINE A SOLID-STATE DEVICE.

DEFINE NEGATIVE TEMPERATURE COEFFICIENT.

NAME THREE OF THE LARGEST USERS OF SEMICONDUCTOR DEVICES.

STATE ONE REQUIREMENT OF AN ELECTRON TUBE, WHICH DOES NOT EXIST FOR SEMICONDUCTORS, THUS MAKING THE ELECTRON TUBE LESS EFFICIENT THAN THE SEMICONDUCTOR.

As you recall from previous studies in this series, semiconductors have electrical properties somewhere between those of insulators and conductors. The use of semiconductor materials in electronic components is not new; some devices are as old as the electron tube. Two of the most widely known semiconductors in use today are the JUNCTION DIODE and TRANSISTOR. These semiconductors fall under a more general heading called solid-state devices. A SOLID-STATE DEVICE is nothing more than an electronic device which operates by the movement of electrons within a solid piece of semiconductor material.

Since the invention of the transistor, solid-state devices have been developed and improved at an unbelievable rate. Great strides have been made in the manufacturing techniques, and there is no foreseeable limit to the future of these devices. Solid-state devices, made from semiconductor materials, offer compactness, efficiency, ruggedness, and versatility. Consequently, these devices have invaded virtually every field of science and industry. In addition to the junction diode and transistor, a whole new family of related devices has been developed which includes the ZENER DIODE, LIGHT-EMITTING DIODE, and FIELD EFFECT TRANSISTOR. One development that has dominated solid-state technology for the last decade and probably has had greater impact on the electronics industry than either the electron tube or transistor is the INTEGRATED CIRCUIT. The integrated circuit is a minute piece of semiconductor material that can produce complete electronic circuit functions.

As the applications of solid-state devices mount, the need for knowledge of these devices becomes increasingly important. Personnel in the Marine Corps today will have to understand solid-state devices if they are to become proficient in the repair and maintenance of electronic equipment. Therefore, the objective of this course is to provide a broad application of power supplies. Our discussion will begin with some background information on the development of the semiconductor. We will then proceed to the semiconductor diode, the transistor, special devices and, finally, solid state power supplies.

Although the semiconductor was late in reaching its present development, its story began long before the electron tube. Historically, we can go far back as 1883 when Michael Faraday discovered that silver sulfide, a semiconductor. has a negative temperature coefficient. The term negative temperature coefficient is just another way of saying its resistance to electrical current flow decreases as temperature increases. The opposite is true of the conductor. It has a positive temperature coefficient. Because of this particular characteristic, semiconductors are used extensively in power-measuring equipment.

Only two years later another valuable characteristic was reported by Munk A. Rosenshold. He found that certain materials have rectifying properties. Strange as it may seem, his finding was given such little notice that it had to be rediscovered 39 years later by F.Braun.

Toward the close of the 19th century, experimenters began to notice the peculiar characteristics of the chemical element SELENIUM. They discovered that in addition to its rectifying properties (the ability to convert a.c. into d.c.), selenium was also light-- its resistance decreases with an increase in light intensity. This discovery eventually led to the invention of the photophone by Alexander Graham Bell.

The photophone, which converted variations of light into sound, was a predecessor of the radio receiver; however, it wasn't until the actual birth of radio that selenium was used to any extent. Today, selenium is an important and widely used semiconductor.

Many other materials were tried and tested for use in communications. SILICON, a metallic element, was found to be the most stable of the materials tested while GALENA, a crystalline form of lead sulfide, was found the most sensitive for use in early radio receivers. By 1915, Carl Beredicks discovered that GERMANIUM, another metallic element, also had rectifying capabilities. Later, it became widely used in electronics for low-power frequency applications.

Although the semiconductor was known long before the electron tube was invented, the semiconductor devices of that time could not match the performance of the tube. Radios needed a device that could not only handle power and amplify but rectify and detect a signal as well. Since tubes could do all these things and the semiconductor devices of that day could not, the semiconductor soon lost out.

It wasn't until the beginning of World War II that interest was renewed in the semiconductor. There was a dire need for a device that could work using the ultra-high frequencies of radar. Electron tubes had interelectrode capacitance that was too high to do the job. The point-contact semiconductor diode, on the other hand, had a very low internal capacitance. Consequently, it fit the bill. It could be designed to work within the ultra-high frequencies used in radar where the electron tube could not.

As radar took on greater importance and communication-electronic equipment became more sophisticated, the demands for better solid-state devices mounted. The limitations of the electron tube made necessary a quest for something new and different. An amplifying device was needed that was smaller, lighter, more deficient, and capable of handling extremely high frequencies. Thus, a serious study of semiconductor materials began in the early 1940's and has continued since.

In June 1948, a significant breakthrough took place in semiconductor development. This was the discovery of the POINT-CONTACT TRANSISTOR. Here, at last, was a semiconductor that could amplify. This discovery brought the semiconductor back into competition with the electron tube. A year later, JUNCTION DIODES AND TRANSISTORS were developed. The junction transistor was found superior to the point-contact type in many respects. By comparison, the junction transistor was more reliable, generated less noise, and had higher power-handling ability than its point-contact brother. The junction transistor became a rival of the electron tube in many uses previously uncontested.

Semiconductor diodes were not to be slighted. The initial work of Dr. Carl Zener led to the development of the ZENER DIODE, which is frequently used today to regulate power supply voltages at precise levels. Considerably more interest in the solid-state diode was generated when Dr. Leo Esaki, a Japanese scientist, fabricated a diode that could amplify. This device, named the TUNNEL DIODE, has amazing gain and fast switching capabilities. Although it is used in the conventional amplifying and oscillating circuits, its primary use is in computer logic circuits.

Another breakthrough came in the late 1950's when it was discovered that semiconductor materials could be combined and treated so that they functioned an entire circuit subassembly rather than as a circuit component. Many names have been given to this solid-circuit concept such as INTEGRATED CIRCUITS, MICROELECTRONICS, and MICROCIRCUITRY.

So we see, in looking back, that the semiconductor is not something new, but it has come a long way in a short time.

In the previous paragraphs, we mentioned just a few of the many different applications of semiconductors devices. The use of these devices has become so widespread that it would be impossible to list all their different applications. Instead, a broad coverage of their specific application is presented.

Semiconductor devices are all around us. They can be found in just about every commercial product we touch, from the family car to the pocket calculator. Semiconductor devices have even found their way into television sets, portable radios, and stereo equipment.

Science and industry also rely on semiconductor devices. Research laboratories use these devices in all sorts of electronic instruments to perform test measurements, and numerous other experimental tasks. Industrial control systems (such as those used to manufacture automobiles) and automatic telephone exchanges also use semiconductors. Heavy duty solid-state diodes are being used to convert large amounts of power for electric railroads. Of the many different applications for solid-state devices, space systems, computers, and data processing equipment are some of the largest consumers.

The various types of modern military equipment are literally loaded with semiconductor devices. Many communication, airborne, and radars are transistorized. Data display systems, data processing units, computers, and aircraft guidance-control assemblies are also good examples of electronic equipments that use semiconductor devices. All of the specific applications of semiconductor devices would make an impressive list. The fact is, semiconductors are now being used extensively in commercial products, industry, and all branches of the Armed Services.

It should not be difficult to conclude, from what you already know, that semiconductor devices can and do perform all the conventional functions of rectification, amplification, oscillation, timing, switching, and sensing. Simply stated, these devices perform the same basic functions as the electron tube but perform more efficiently, economically, and for a longer period of time. Therefore, it should be no surprise to you to see these devices used in place of electron tubes. Keeping this in mind, we see that it is only natural and logical to compare semiconductor devices with electron tubes.

Physically, semiconductor devices are much smaller than tubes. You can see in figure 1-1 that the difference is quite evident. This illustration shows some commonly used tube sizes alongside semiconductor devices of similar capabilities. The reduction in size can be as great as 100:1 by weight and 1000:1 by volume.

It is easy to see that size reduction favors the semiconductor device. Therefore, whenever miniaturization is required or is convenient, transistors are favored over tubes. Bear in mind, however, that the extent of practical size reduction is a big factor,and many things must be considered. Miniature electron tubes, for example, may be preferred in certain applications to transistors, thus keeping size reduction a competitive area.

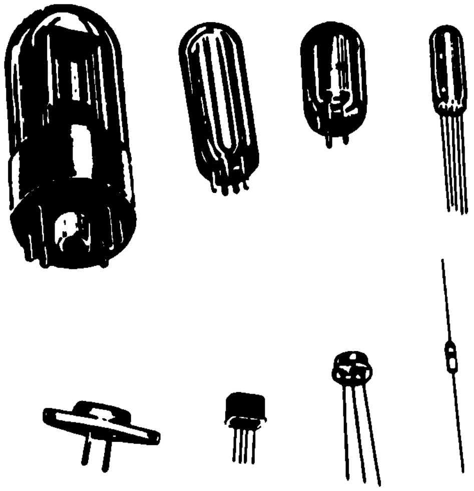


Fig 1-1. Size comparisons of electron tubes and semiconductors.

Power is also a two-sided story. For low-power applications, where efficiency is a significant factor, semiconductors have a decided advantage. This is true mainly because semiconductor devices perform very well with an extremely small amount of power. In addition, they require no filaments of heaters as in the case of the electron tube. For example, a computer operating with over 4000 solid-state devices may require no more than 20 watts of power. However, the same number of tubes would require several kilowatts of power.

For high-power applications, it is a different story--tubes have the upper hand. The high-power electron tube has no equivalent in any semiconductor device. This is because a tube can be designed to operate with over a thousand volts applied to its plate; whereas, the maximum allowable voltage for a transistor is limited to about 200 volts (usually 50 volts or less). A tube can also handle thousands of watts of power. The maximum power output for transistors generally ranges from 30 milliwatts to slightly over 100 watts.

When it comes to ruggedness and life expectancy, the tube is still in the competition. Design and functional requirements usually dictate the choice of devices. However, semiconductor devices are rugged and long-lived. They can be constructed to withstand extreme vibration and mechanical shock. They have been known to withstand impacts that would completely shatter an ordinary electron tube. Although some specially designed tubes render extensive service, the life expectancy of transistors is better than three to four times that of ordinary electron tubes. There is no know failure mechanism (such as an open filament in a tube) to limit the semiconductor's life. However, semiconductor devices do have some limitations. They are usually affected more by temperature, humidity, and radiation than are tubes.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is a solid-state device?

2. Define negative temperature coefficient.

3. Name three of the largest users of semiconductor devices.

   a.

   b.

   c.

4. State one requirement of an electron tube, which does not exist for semiconductors, that makes the tube less efficient than the semiconductor.

Work Unit 1-2. SEMICONDUCTOR THEORY

DEFINE MATTER AND LIST ITS THREE DIFFERENT STATES.

NAME THE OUTER SHELL OF AN ATOM.

STATE, IN TERMS OF ENERGY BANDS, WHETHER A SUBSTANCE IS A GOOD INSULATOR, SEMICONDUCTOR, OR CONDUCTOR.

NAME THE TERM USED TO DESCRIBE THE CLOSING OF VALANCE ELECTRONS BETWEEN TWO OR MORE ATOMS.

NAME THE TWO TYPES OF CURRENT FLOW IN A SEMICONDUCTOR.

STATE THE NAME GIVEN TO A DOPED GERMANIUM CRYSTAL WITH AN EXCESS OF FIVE HOLES.

To understand why solid-state devices function as they do, we will have to examine closely the composition and nature of semiconductors. This entails theory which is fundamental to the study of solid-state devices. Rather than beginning with theory, lets first become reacquainted with some of the basic information concerning matter and energy.

The universe, as we know it today, is divided into two parts, matter and energy. Matter, which is our main concern at this time, is anything that occupies space and has weight. Rocks, water, air, automobiles, clothing, and even our own bodies are good examples of matter. From this, we can conclude that matter may be found in any one of three states: SOLID, LIQUID, and GASEOUS. All matter is composed of either an element or combination of elements. As you know, an element is a substance which cannot be reduced to a simpler form by chemical means. Examples of elements with which you are in contact with everyday are iron, gold, silver, copper, and oxygen. At present, there are over 100 known elements of which matter is comprised.

As we work our way down the size scale, we come to the atom, the smallest particle into which an element can be broken down and still retain all its original properties. The atom of one element, however, differs from the atoms of all other elements. Since there are over 100 known elements, there must be over 100 different atoms, or a different atom for each element.

Now let us consider more than one element at a time. This brings us to the term, "compound." A compound is a chemical combination of two or more elements. Water, table salt, ethyl alcohol, and ammonia are all examples of compounds. The smallest part of a compound, which has all the characteristics of the compound, is the molecule. Each molecule contains some of the atoms of each of the elements forming the compound.

Consider sugar, for example. Sugar in general terms is matter, since it occupies space and has weight. It is also a compound because it consists of two or more elements. Take a lump of sugar and crush it into small particles; each of the particles still retains its original identifying properties of sugar. The only thing that changed was the physical size of the sugar. If we continue this subdividing process by grinding the sugar into a fine powder, the results are the same. Even dissolving the sugar in water does not change its identifying properties, in spite of the fact that the particles of sugar are now too small to see even with a microscope. Eventually, we end up with a quantity of sugar which cannot be further divided without its ceasing to be sugar. This quantity is known as a molecule of sugar. If the molecule is further divided, it is found to consist of three simpler kinds of matter: carbon, hydrogen, and oxygen. These simpler forms are called elements. Therefore, since elements consist of atoms, then a molecule of sugar is made up of atoms of carbon, hydrogen, and oxygen.

As we investigate the atom, we find that it is basically composed of electrons, protons, and neutrons. Furthermore, the electrons, protons, and neutrons of one element are identical to those of any other element. There are different kinds of elements because the number and the arrangement of electrons and protons are different for each element.

The electron is considered to be a small negative charge of electricity. The proton has a positive charge of electricity equal and opposite to the charge of the electron. Scientists have measured the mass and size of the electron and the proton; therefore, it is known how much charge each possesses. Both the electron and proton have the same quantity of charge, although the mass of the proton is approximately 1837 times that of the electron. In some atoms there exists a neutral particle called a neutron. The neutron has a mass approximately equal to that of a proton, but it has no electrical charge.

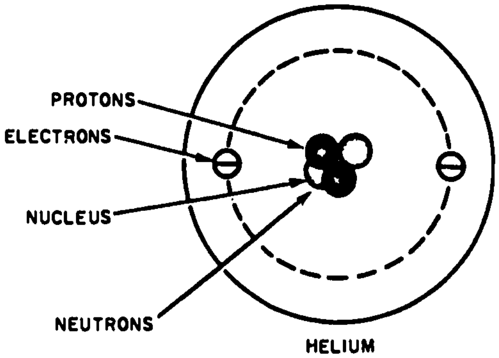


Fig 1-2. The composition of a simple helium atom.

According to a popular theory, the electrons, protons, and neutrons of the atoms are thought to be arranged in a manner similar to a miniature solar system. Notice the helium atom in figure 1-2. Two protons and two neutrons form the heavy nucleus with a positive charge around which two very light electrons revolve. The path each electron takes around the nucleus is called an orbit. The electrons are continuously being acted upon in their orbits by the force of attraction of the nucleus. To maintain an orbit around the nucleus, the electrons travel at a speed that produces a counterforce equal to the attraction force of the nucleus. Just as energy is required to move a space vehicle away from the earth, energy is also required to move an electron away from the nucleus. Like a space vehicle, the electron is said to be at a higher energy level when it travels a larger orbit. Scientific experiments have shown that the electron requires a certain amount of energy to stay in orbit. This quantity is called the electron's energy level. By virtue of just its motion alone, the electron contains kinetic energy. Due to its position, it also contains potential energy. The total energy contained by an electron (kinetic energy plus potential energy) is the main factor which determines the radius of the electron's orbit. In order for an electron to remain in this orbit, it must neither gain or lose energy.

The orbiting electrons do not follow random paths, instead they are confined to definite energy levels. Visualize these levels as shells with each successive shell being spaced a greater distance from the nucleus. The shells and the number of electrons required to fill them, may be predicted by using Pauli's exclusion principle. Simply stated, this principle specifies that each shell will contain a maximum of $2n^2$ electrons, where it corresponds to the shell number starting with the one closest to the nucleus. By this principle, the second shell, for example, would contain $2(2)^2$ or 8 electrons when full.

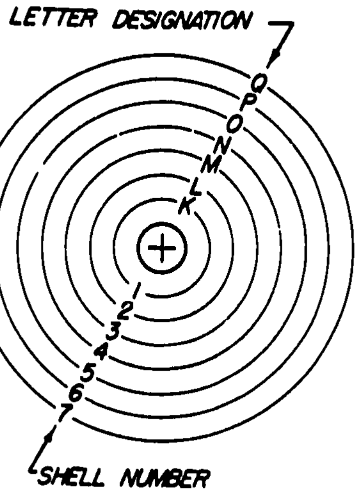


Fig 1-3. Shell designation.

In addition to being numbered, the shells are also given letter designations starting with the shell closest to the nucleus and progressing outward as shown in figure 1-3. The shells are considered to be full, or complete, when they contain the following quantities of electrons: two in the K(1st) shell, eight in the L(2nd) shell, eighteen in the M(3rd) shell, and so on, in accordance with the exclusion principle. Each of these shells is a major shell and can be divided into subshells, of which there are four, labeled s, p, d, and f. Like the major shells, the subshells are also limited as to the number of electrons which they contain. Thus, the "s" subshell is complete when it contains two electrons, the "p" subshell when it contains six, the "d" subshell when it contains ten, and the "f" subshell when it contains fourteen electrons.

In as much as the K shell can contain no more than two electrons, it must have only one subshell, the s subshell. The M shell is composed of three subshells: s, p, and d. If the electrons in the s, p, and d subshells are added together, their total is found to be 18, the exact number required to fill the M shell. Notice the electron configuration for copper illustrated in figure 1-4. The copper atom contains 29 electrons, which completely fill the first three shells and subshells, leaving one electron in the "s" subshell on the N shell. A list of all the other known elements, with the number of electrons in each atom, is contained in the PERIODIC TABLE OF ELEMENTS. This table is located on page 1-30.

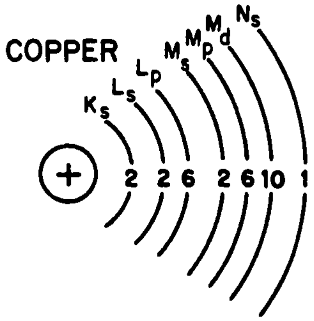


Fig 1-4. Copper atom.

Valence is an atom's ability to combine with other atoms. The number of electrons in the outermost shell of an atom determines its valence. For this reason, the outer shell of an atom is called the VALENCE SHELL, and the electrons contained in this shell are called VALENCE ELECTRONS. The valence of an atom determines its ability to gain or lose an electron, which, in turn, determines the chemical and electrical properties of the atom. An atom that is lacking only one or two electrons from its outer shell will easily gain electrons to complete its shell, but a large amount of energy is required to free any of its electrons. An atom having a relatively small number of electrons in its outer shell in comparison to the number of electrons required to fill the shell will easily lose these valence electrons. The valence shell always refers to the outermost shell.

Now that you have become reacquainted with matter and energy, we will continue our discussion with electron behavior.

As stated earlier, orbiting electrons contain energy and are confined to definite energy levels. The various shells in an atom represent these energy levels. Therefore, in order to move an electron from a lower shell to a higher shell, a certain amount of energy is required. This energy can be in the form of electric fields, heat, light, and even bombardment by other particles. Failure to provide enough energy to the electron, even if the energy supplied is just short of the required amount, will cause it to remain at its present energy level. Supplying more energy than is needed will only cause the electron to move to the next higher shell and the remaining energy will be wasted. In simple terms, energy is required in definite units to move electrons from one shell to the next higher shell. These units are called QUANTA (for example, 1, 2, or 3 quanta).

Electrons can also lose energy as well as receive it. When an electron loses energy, it moves to a lower shell. The lost energy, in some cases, appears as heat.

If a sufficient amount of energy is absorbed by an electron, it is possible for that electron to be completely removed from the influence of the atom. This is called IONIZATION. When an atom loses electrons or gains electrons in this process of electron exchange, it is said to be ionized. For ionization to take place there must be a transfer of energy which results in a change in the internal energy of the atom. An atom having more than its normal amount of electrons acquires a negative charge and is called a NEGATIVE ION. The atom that gives up some of its normal electrons is left with fewer negative charges than positive and is called a POSITIVE ION. Thus, we can define ionization as the process by which an atom loses or gains electrons.

Up to this point in our discussion we have spoken only of isolated atoms. When atoms are spaced far enough apart, as in a gas, they have very little influence upon each other, and are very much like lone atoms, but atoms within a solid have a marked effect upon each other. The forces that bind these atoms together greatly modify the behavior of the other electrons. One consequence of this close proximity of atoms is to cause the individual energy levels of an atom to break up and form bands of energy. Discrete (separate and complete) energy levels still exist within these energy bands, but there are many more energy levels than there were with the isolated atom. In some cases, energy levels will have disappeared. Figure 1-5 shows the difference in the energy arrangement between an isolated atom and the atom in a solid. Notice that the isolated atom (such as in gas) has energy levels; whereas, the atom in a solid has energy levels grouped into ENERGY BANDS.

The upper band in the solid in figure 1-5 is called the CONDUCTION BAND because electrons in this band are easily removed by the application of external electric fields. Materials that have a large number of electrons in the conduction band act as good conductors of electricity.

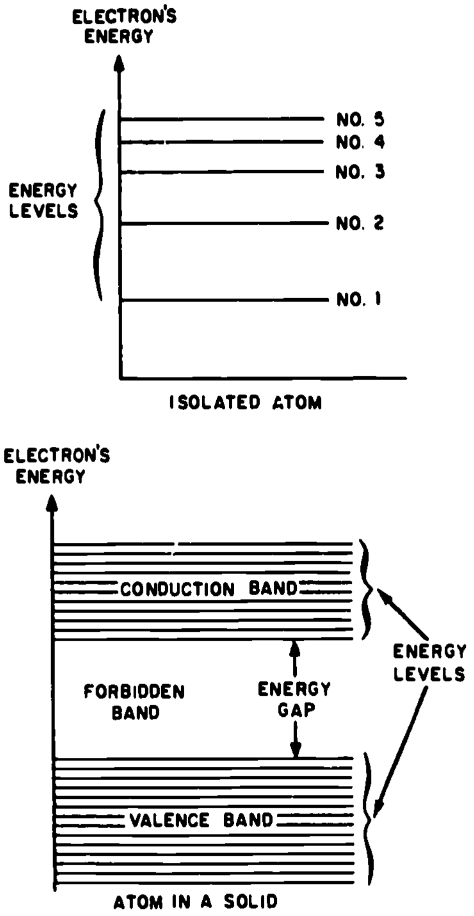


Fig 1-5. The energy arrangement in atoms.

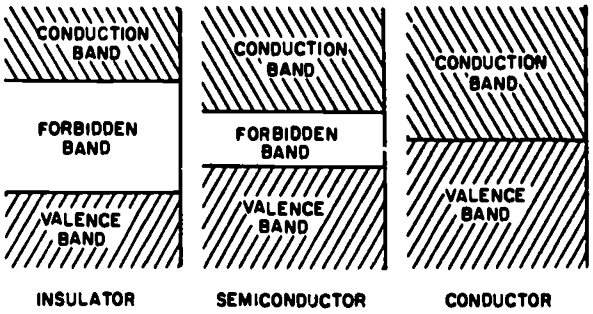


Fig 1-6. Energy level diagrams.

Below the conduction band is the FORBIDDEN BAND or energy gap. Electrons are never found in this band, but may travel back and forth through it, provided they do not come to rest in the band.

The last band or VALENCE BAND is composed of a series of energy levels containing valence electrons. Electrons in this band are more tightly bound to the individual atom than the electrons in the conduction band. However, the electrons in the valence band can still be moved to the conduction band with the application of energy, usually thermal energy. There are more bands below the valence band, but they are not important to the understanding of semiconductor theory and will, therefore, not be discussed.

The concept of energy bands is particularly important in classifying materials as conductors, semiconductors, and insulators. An electron can exist in either of two energy bands, the conduction band or the valence band. All that is necessary to move an electron from the valence band to the conduction band, so it can be used for electric current, is enough energy to carry the electron through the forbidden band. The width of the forbidden band or the separation between the conduction and valence bands determines whether a substance is an insulator, semiconductor, or conductor. Figure 1-6 uses energy level diagrams to show the difference between insulators, semiconductors, and conductors.

The energy diagram for the insulator shows the insulator with a very wide energy gap. The wider this gap, the greater the amount of energy required to move the electron from the valence band to the conduction band. Therefore, an insulator requires a large amount of energy to obtain a small amount of current. The insulator "insulates" because of the wide forbidden band or energy gap.

The semiconductor, on the other hand, has a smaller forbidden band and requires less energy to move an electron from the valence band to the conduction band. Hence, for a certain amount of applied voltage, more current will flow in the semiconductor than in the insulator.

The last energy level diagram in figure 1-6 is that of a conductor. Notice, there is no forbidden band or energy gap and the valence and conduction bands overlap. With no energy gap, it only takes a small amount of energy to move electrons into the conduction band; consequently, conductors pass electrons very easily.

The chemical activity of an atom is determined by the number of electrons in its valence shell. When the valence shell is complete, the atom is stable and shows little tendency to combine with other atoms to form solids. Only atoms that possess eight valence electrons have a complete outer shell. These atoms are referred to as inert or inactive atoms. However, if the valence shell of an atom is short the required number of electrons to complete the shell, then the activity of the atom increases.

Silicon and germanium, for example, are the most frequently used semiconductors. Both are quite similar in their structure and chemical behavior. Each has four electrons in the valence shell. Consider just silicon. Since it has fewer than the required number of eight electrons needed in the outer shell, its atoms will unite with other atoms until eight electrons are shared. This gives each atom a total of eight electrons in its valence shell—four of its own and four that it borrowed from the surrounding atoms. The sharing of valence electrons between two or more atoms produces a COVALENT BOND between the atoms. It is this bond that holds the atoms together in an orderly structure called a CRYSTAL. A crystal is just another name for a solid whose atoms or molecules are arranged in a three-dimensional geometrical pattern commonly referred to as a lattice. Figure 1-7 shows a typical crystal structure. Each sphere in the figure represents the nucleus of an atom, and the arms that join the atoms and support the structure are the covalent bonds.

As a result of this sharing process, the valence electrons are held tightly together. This can be best illustrated by the two-dimensional view of the silicon lattice in figure 1-8. The circles in the figure represent the nuclei of the atoms. The +4 in the circles is the net charge of the nucleus plus the inner shells (minus the valence shell). The short lines indicate valence electrons. Because every atom in this pattern is bonded to four other atoms, the electrons are not free to move within the crystal. As a result of this bonding pure silicon and germanium are poor conductors of electricity. The reason they are not insulators but semiconductors is because with the proper application of heat or electrical pressure, electrons can be caused to break free of their bonds and move into the conduction band. Once in this band, they wander aimlessly through the crystal.

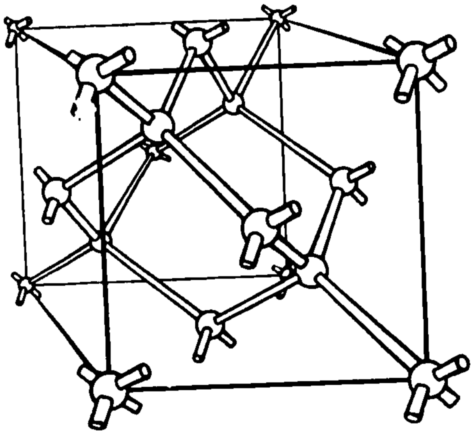


Fig 1-7. A typical crystal structure.

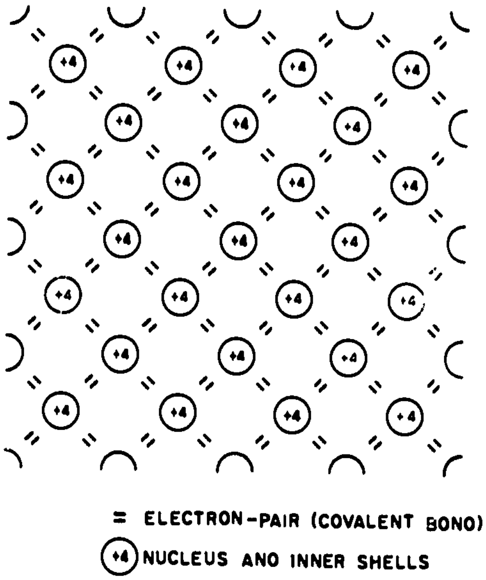


Fig 1-8. A two dimensional view of a silicon cubic lattice.

As stated earlier, energy can be added to electrons by applying heat. When enough energy is absorbed by the valence electrons, it is possible for them to break some of their covalent bonds. Once the bonds are broken, the electrons move to the conduction band where they are capable of supporting electric current. When a voltage is applied to a crystal containing these conduction band electrons, the electrons move through the crystal toward the applied voltage. This movement of electrons in a semiconductor is referred to as electron current flow.

There is still another type of current in a pure semiconductor. This current occurs when a covalent bond is broken and a vacancy is left in the atom by the missing valence electron. This vacancy is commonly referred to as a "hole." The hole is considered to have a positive charge because its atom is deficient by one electron which causes the protons to outnumber the electrons. As a result of this hole, a chain reaction begins when a nearby electron breaks its own covalent bond to fill the hole, leaving another hole. Then another electron breaks its bond to fill the previous hole, leaving still another hole. Each time an electron in this process fills a hole, it enters into a covalent bond. Even though an electron has moved from one covalent bond to another, the most important thing to remember is that the hole is also moving. Therefore, since this process of conduction resembles the movement of holes rather than electrons, it is termed hole flow (short for hole current flow or conduction by holes). Hole flow is very similar to electron flow except that the holes move toward a negative potential and in an opposite direction to that of the electron. Since hole flow results from the breaking of covalent bonds, which are at the valence band level, then the electrons associated with this type of conduction contain only valence band energy and must remain in the valence band. However, the electrons associated with electron flow have conduction band energy and can, therefore, move throughout the crystal. A good analogy of hole flow is the movement of a hole through a tube filled with balls (fig 1-9).

When ball number 1 is removed from the tube, a hole is left. This hole is then filled by ball number 2, which leaves still another hole. Ball number 3 then moves into the hole left by ball number 2. This causes still another hole to appear where ball 3 was located. Notice the holes are moving to the right side of the tube. This action continues until all the balls have moved one space to the left in which time the hole moved eight spaces to the right and came to rest at the right-hand end of the tube.

In the theory just described, two-current carries were created by the breaking of covalent bonds: the negative electron and the positive hole. These carriers are referred to as electron-hole pairs. Since the semiconductor we have been discussing contains no impurities, the number of holes, in the electron-hole pairs, is always equal to the number of conduction electrons. Another way of describing this condition where no impurities exist is by saying the semiconductor is INTRINSIC. The term intrinsic is also used to distinguish the pure semiconductor that we have been working with from one containing impurities.

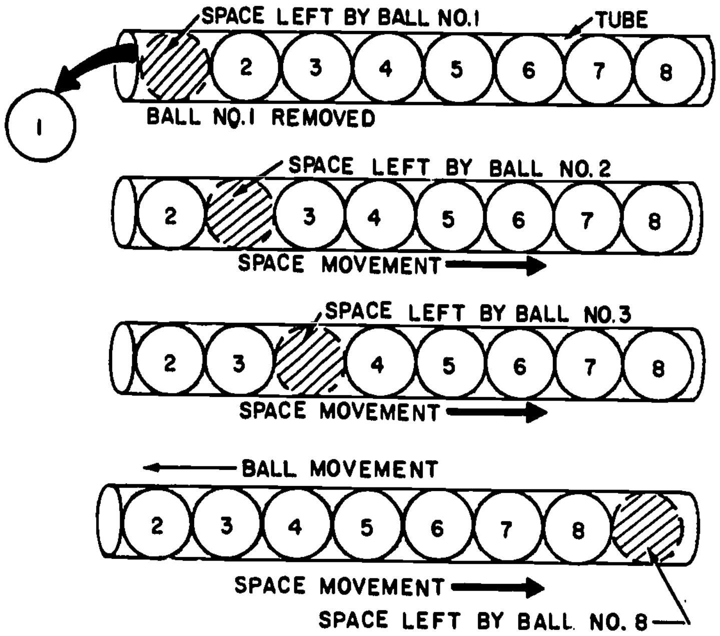


Fig 1-9. Analogy of hole flow.

The pure semiconductor mentioned earlier is basically neutral. It contains no free electrons in its conduction bands. Even with the application of thermal energy, only a few covalent bonds are broken yielding a relatively small current flow. A much more efficient method of increasing current flow in semiconductors is by adding very small amounts of selected additives to them, generally no more than a few parts per million. These additives are called impurities and the process of adding them to crystals is referred to as DOPING. The purpose of semiconductor doping is to increase the number of free charges that can be moved by an external applied voltage. When an impurity increases the number of free electrons, the doped semiconductor is NEGATIVE or N TYPE, and the impurity that is added is known as an N-type impurity.

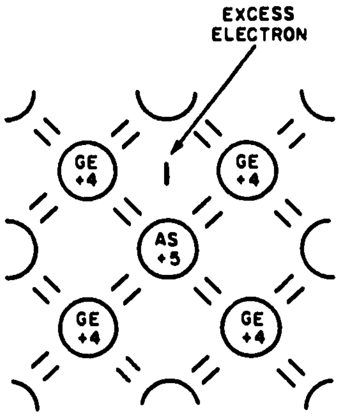


Fig 1-10. Germanium crystal doped with arsenic.

However, an impurity that reduces the number of free electrons, causing more holes, creates a POSITIVE or P-TYPE semiconductor, and the impurity that was added to it is known as a P-type impurity. Semiconductors which are doped in this manner, either with N- or P-type impurities, are referred to as EXTRINSIC semiconductors.

N-Type Semiconductor

The N-type impurity loses its extra valence electron easily when added to a semiconductor material, and in so doing increases the conductivity of the material by contributing a free electron. This type of impurity has 5 valence electrons and is called a PENTAVALENT impurity. Arsenic, antimony, bismuth, and phosphorous are pentavalent impurities. Because these materials give or donate one electron to the doped material, they are also called DONOR IMPURITIES.

When a pentavalent (donor) impurity, like arsenic, is added to germanium, it will form covalent bonds with the germanium atoms. Figure 1-10 illustrates this by showing an arsenic atom (AS) in a germanium (GE) lattice structure. Notice this arsenic atom in the center of the lattice. It has 5 valence electrons in its outer shell but uses only 4 of them to form covalent bonds with the germanium atoms, leaving one electron relatively free in the crystal structure. Pure germanium may be converted into a N-type semiconductor by "doping" it with any donor impurity having five valence electrons in its outer shell. Since this type of semiconductor (N-type) has a surplus of electrons, the electrons are considered MAJORITY carriers, while the holes, being few in number, are the MINORITY carriers.

P-Type Semiconductor

The second type of impurity when added to a semiconductor material tends to compensate for its deficiency of 1 valence electron by acquiring an electron from its neighbor. Impurities of this type have only 3 valence electrons and are called TRIVALENT impurities. Aluminum, indium, gallium, and boron are trivalent impurities. Because these materials accept 1 electron from the doped material they are also called ACCEPTOR impurities.

A trivalent (acceptor) impurity element can also be used to dope germanium. In this case, the impurity is 1 electron short of the required amount of electrons needed to establish covalent bonds with 4 neighboring atoms. Thus, in a single covalent bond there will be only 1 electron instead of 2. This arrangement leaves a hole in that covalent bond. Figure 1-11 illustrates this theory by showing what happens when germanium is doped with an indium (In) atom. Notice, the indium atom in the figure is 1 electron short of the required amount of electrons needed to form a covalent bonds with 4 neighboring atoms and, therefore, creates a hole in the structure.

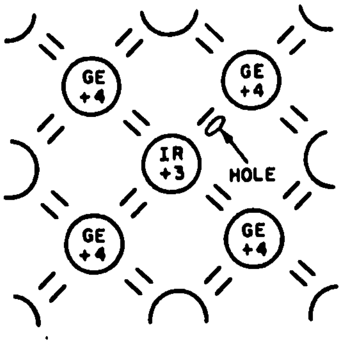


Fig 1-11. Germanium crystal doped with indium.

Gallium and boron, which are also trivalent impurities, exhibit these same characteristics when added to germanium. The holes can only be present in this type semiconductor when a trivalent impurity is used. Note that a hole carrier is not created by the removal of an electron from a neutral atom, but is created when a trivalent impurity enters into covalent bonds with a tetravalent (4 valence electrons) crystal structure. The holes in the type semiconductor (P-type) are considered the MAJORITY carries since they are present in the material in the greatest quantity. The electrons, on the other hand, are the MINORITY carries.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. Define matter.

2. List the three different states of matter.

   a. __________

   b. __________

   c. __________

3. What is the outer shell of an atom called?

4. State, in terms of energy bands, whether a substance is a good insulator, semiconductor, or conductor.

5. What term is used to describe the definite discrete amounts of energy required to move an electron from a lower shell to a higher shell?

6. What is the term used to describe the sharing of valence electrons between two or more atoms?

7. State the two types of current flow in a semiconductor.

   a. ____

   b. ____

8. What is the name given to a piece of pure semiconductor material that has an equal number of electrons and holes?

   ____

9. What is the name given to a doped germanium crystal with an excess of free holes?

   ____

10. What are the majority of carriers in an N-Type semiconductor?

   ____

Work Unit 1-3. SEMICONDUCTOR DIODE

STATE THE PURPOSE OF A PN JUNCTION DIODE.

SPECIFY, IN REFERENCE TO THE SCHEMATIC SYMBOL FOR A DIODE, THE DIRECTION OF ELECTRON FLOW.

SPECIFY, IN ORDER TO REVERSE BIAS IN A PN JUNCTION, THE TERMINAL OF A BATTERY THAT IS CONNECTED TO THE P MATERIAL.

STATE THE TYPE OF BIAS WHICH OPPOSES THE PN JUNCTION BARRIER.

If we join a section of N-type semiconductor material with a similar section of P-type semiconductor material, we obtain a device known as a <u>PN JUNCTION</u>. (The area where the N and P regions meet is appropriately called the junction.) The unusual characteristics of this device make it extremely useful in electronics as a diode rectifier. The diode rectifier or PN junction diode performs the same functions as its counterpart in electron tubes but in a different way. The diode is nothing more than a two-element semiconductor device that makes use of the rectifying properties of a PN junction to convert alternating current into direct current by permitting current flow in only one direction. The schematic symbol of a PN junction diode is shown in figure 1-12. The vertical bar represents the <u>cathode</u> (N-type material) since it is the source of electrons and the arrow represents the <u>anode</u> (P-type material) since it is the destination of the electrons. The label "CR1" is an alphanumerical code used to identify the diode. In this figure, we have only one diode so it is labeled CRI (crystal rectifier number one). If there were four diodes shown in the diagram then the last diode would be labeled CR4. The heavy dark line shows electron flow. Notice it is against the arrow. For further clarification, a pictorial diagram for a PN junction and an actual semiconductor (one of many types) are also illustrated.

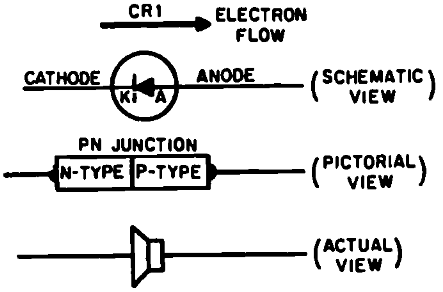


Fig 1-12. The PN junction diode.

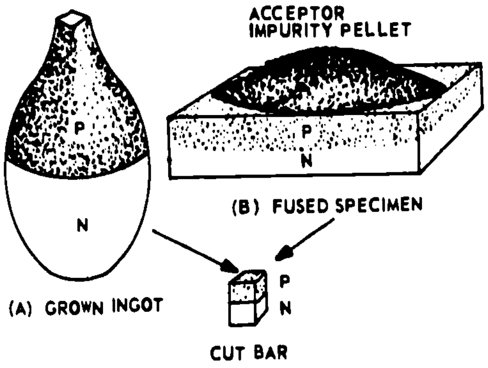


Fig 1-13. Grown and fused PN junctions from which bars are cut.

Merely pressing together a section of P material and a section of N material, however, is not sufficient to produce a rectifying junction. To form a proper PN junction, the semiconductor should be in one piece, but divided into a P-type impurity region and an N-type impurity region. This can be done in various ways. One way is to mix P-type and N-type impurities into a single crystal during the manufacturing process. By so doing, a P-region is grown over part of a semiconductor's length and an N-region is grown over the other part. This is called a GROWN junction and is illustrated in view (A) of figure 1-13. Another way to produce a PN junction is to melt one typical type of impurity into a semiconductor of the opposite type impurity. For example, a pellet of acceptor impurity is placed on a wafer of N-type germanium and heated. Under controlled temperature conditions, the acceptor impurity fuses into the wafer to form a P-region within it, as shown in view (B) of figure 1-13. This type of junction is known as an ALLOY or FUSED-ALLOY junction, and it is one of the most commonly used junctions. In figure 1-14, a POINT-CONTACT type of construction is shown. It consists of a fine metal wire, called a cat whisker, that makes contact with a small area on the surface of an N-type semiconductor as shown in view (A) of the figure. The PN union is formed in this process by momentarily applying a high-surge current to the wire and the N-type semiconductor. The heat generated by this current converts the material nearest the point of contact to a P-type material (view B).

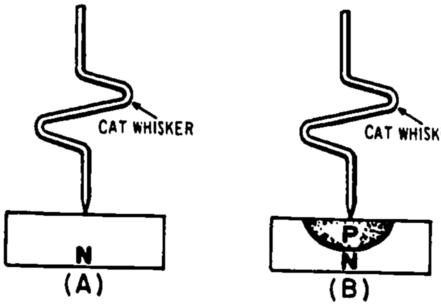


Fig 1-14. The point-contact type of diode construction.

Still another process is to heat a section of semiconductor material to near melting and then diffuse the impurity atoms into a surface layer. Regardless of the process, the object is to have a perfect bond everywhere along the union (interface) between the P and N materials. Proper contact along the union is important because, as we will see later, the union (junction or interface) is the rectifying agent in the diode.

Now that you are familiar with P- and N-type materials, how these materials are joined together to form a diode, and the function of the diode, let us continue our discussion with the operation of the PN junction. However, before you can understand how the PN junction works, we must first consider current flow in the materials that make up the junction and then what happens initially within the junction when these two materials are joined together.

Conduction in the N-type of semiconductor, or crystal, is similar to conduction in a copper wire. That is, with voltage applied across the material, electrons will move through the crystal just as current would flow in a copper wire. This is shown in figure 1-15. The positive potential of the battery will attract the free electrons in the crystal. These electrons will leave the crystal and flow into the positive terminal of the battery. As an electron leaves the crystal, an electron from the negative terminal of the battery will enter the crystal, thus completing the current path. Therefore, the majority current carries in the N-type material (electrons) are repelled by the negative side of the battery and move through the crystal toward the positive side of the battery.

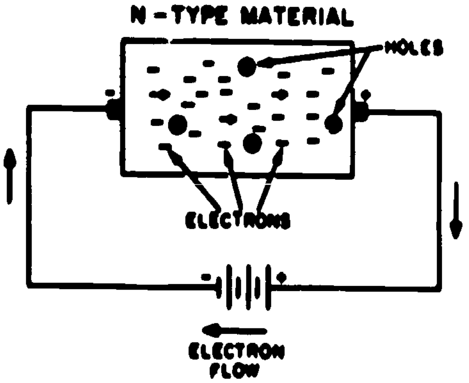


Fig 1-15. Current flow in the N-type material.

Current flow through the P-type material is illustrated in figure 1-16. Conduction in the P material is by positive holes, instead of negative electrons. The hole moves from the positive terminal of the P material to the negative terminal. Electrons from the external circuit enter the negative terminal of the material and fill holes in the vicinity of this terminal. At the positive terminal, electrons are removed from the covalent bonds, thus creating new holes. This process continues as the steady stream of holes (hole current) moves toward the negative terminal.

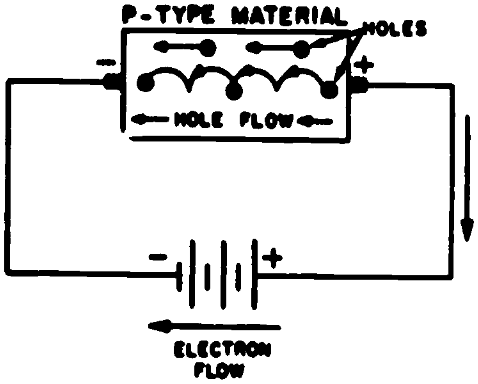


Fig 1-16. Current flow in the P-type material.

Notice in both N-type and P-type materials, current flow in the external circuit consists of electrons moving out the negative terminal of the battery and into the positive terminal of the battery. Hole flow, on the other hand, only exists within the material itself.

Although the N-type material has an excess of free electrons, its still electrically neutral. This is because the donor atoms in the N material were left with positive charges after free electrons became available by covalent bonding (the protons outnumbered the electrons). Therefore, for every free electron in the N material there is a corresponding positively charged atom to balance it. The end result is that the N material has an overall charge of zero.

By the same reasoning, the P-type material is also electrically neutral because the excess of holes in this material is exactly balanced by the number of electrons. Keep in mind that the holes and electrons are still free to move in the material because they are only loosely bound to their parent atoms.

It would seem that if we joined the N and P materials together by one of the processes mentioned earlier that all the holes and electrons would pair up. On the contrary, this does not happen. Instead the electrons in the N material diffuse (move or spread out) across the junction into the P material and fill some of the holes. At the same time the holes in the P material diffuse across the junction into the N material and are filled by N material electrons. This process is called JUNCTION RECOMBINATION reducing the number of free electrons and holes in the vicinity of the junction. Because there is a depletion, or lack of free electrons and holes in this area, it is known as the DEPLETION REGION.

The loss of an electron from the N-type material created a positive ion in the N material while the loss of a hole from the P material created a negative ion in the material. These ions are fixed in place in the crystal lattice structure and cannot move. Thus, they make up a layer of fixed charges on the two sides of the junction as shown in figure 1-17. On the N side of the junction, there is a layer of positively charged ions; on the P side of the junction there is a layer of negatively charged ions. An electrostatic field, represented by a small battery in the figure, is established across the junction between the oppositely charged ions. The diffusion of electrons and holes across the junction will continue until the magnitude of the electrostatic field is increased to the point where the electrons and holes no longer have enough energy to overcome it, and are repelled by the negative and positive ions respectively. At this point equilibrium is established and, for all practical purposes, the movement of carriers across the junction cease. For this reason, the electrostatic field created by the positive and negative ions in the depletion region is called a barrier.

The action just described occurs almost instantly when the junction is formed. Only the carriers in the immediate vicinity of the junction are affected. The carriers throughout the remainder of the N and P material are relatively undisturbed and remain in a balanced condition.

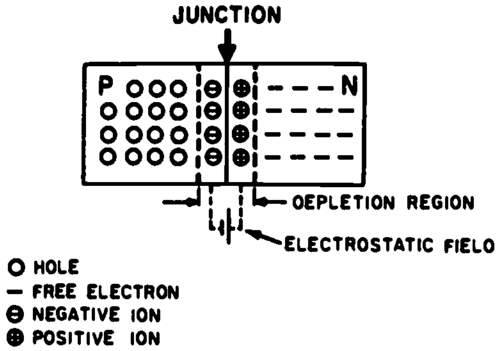


Fig 1-17. The PN junction barrier formation.

FORWARD BIAS-an external voltage applied to a PN junction is called BIAS. If, for example, a battery is used to supply bias to a PN junction and is connected so that its voltage opposes the junction field, it will reduce the junction barrier and therefore aid current flow through the junction. This type of bias is known as FORWARD BIAS, and it causes the junction to offer only minimum resistance to the flow of current.

Forward bias is illustrated in figure 1-18. Notice the positive terminal of the bias battery is connected to the P-type material and the negative terminal of the battery is connected to the N-type material. The positive potential repels holes toward the junction where they neutralize some of the negative ions. At the same time, the negative potential repels electrons toward the junction where they neutralize some of the positive ions. Since ions on both sides of the barrier are being neutralized, the width of the barrier decreases. Thus, the effect of the battery voltage in the forward-bias direction is to reduce the barrier potential across the junction to allow majority carriers to cross the junction. Current flow in the forward-biased PN junction is relatively simple. An electron leaves the negative terminal of the battery and moves to the terminal of the N-type material. It enters the N material, where it is the majority carrier, and moves to the edge of the junction barrier.

Due to forward bias, the barrier offers less opposition to the electron and it will pass through the depletion region into the P-type material. The electron loses energy in overcoming the opposition of the junction barrier, and upon entering the P material, combines with a hole. The hole was produced when an electron was extracted from the P material by the positive potential of the battery. The created hole moves through the P material toward the junction where it combines with an electron.

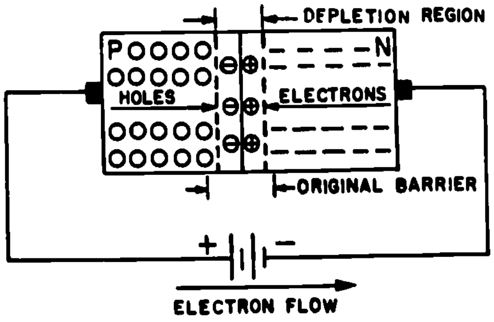


Fig 1-18. Forward-biased PN junction.

It is important to remember that in the forward biased condition, conduction is by MAJORITY current carriers (holes in the P-type material and electrons in the N-type material). Increasing the battery voltage will increase the number of majority carriers arriving at the junction and will, therefore, increase the current flow. If the battery voltage is increased to the point where the barrier is greatly reduced, a heavy current will flow and the junction may be damaged from the resulting heat.

REVERSE BIAS. If the battery mentioned earlier is connected across the junction so that its voltage aids the junction, it will increase the junction barrier and thereby offer a high resistance to the correct flow through the junction. This type of bias is known as reverse bias.

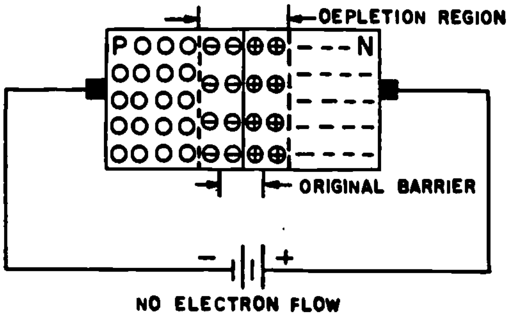


Fig 1-19. Reverse biased PN junction.

To reverse bias a junction diode, the negative battery terminal is connected to the P-type material, and the positive battery terminal to the N-type material as shown in figure 1-19. The negative potential attracts the holes away from the edge of the junction barrier on the P side, while the positive potential attracts the electrons away from the edge of the barrier on the N side. This action increases the barrier width because there are more negative ions on the P side of the junction, and more positive ions on the N side of the junction. Notice in the figure the width of the barrier has increased. The increase in the number of ions prevents current flow across the junction by majority carriers. However, the current flow across the barrier is not quite zero because of the minority carriers crossing the junction. As you recall, when the crystal is subjected to an external source of energy (light, heat, etc.), electron-hole pairs are generated. The electron-hole pairs produce minority current carriers. There are minority current carriers in both regions: holes in the N material and electrons in the P material. With reverse bias, the electrons in the P-type material are repelled toward the junction by the negative terminal of the battery. As the electron moves across the junction, it will neutralize a positive ion in the N-type material. Similarly, the holes in the N-type material will be repelled by the positive terminal of the battery toward the junction. As the hole crosses the junction, it will neutralize a negative ion in the P-type material. This movement of minority carriers is called MINORITY CURRENT FLOW, because the holes and electrons involved come from the electron-hole pairs that are generated in the crystal lattice structure, and not from the addition of impurity atoms.

Therefore, when a PN junction is reverse biased, there will be no current flow due to majority carriers but a very small amount of current due to minority carriers crossing the junction. However, at normal operating temperatures this small current may be neglected.

In summary, the most important point to remember about the PN junction diode is its ability to offer very little resistance to current flow in the forward-bias direction but maximum resistance to current flow when reverse biased.

A good way of illustrating this point is by plotting a graph of the applied voltage versus the measured current. Figure 1-20 shows a plot of this voltage-current relationship (characteristic curve) for a typical PN junction diode.

To determine the resistance from the curve in this figure we can use Ohm's law, $R = \frac{E}{I}$. For example at point A the forward-bias voltage is 1 volt and the forward-bias current is 5 milliamperes. This represents 200 ohms of resistance (1 volt/5 mA = 200 ohms). However, at point B the voltage is 3 volts and the current is 500 milliamperes. This results in 60 ohms of resistance of the diode. Notice that when the forward-bias voltage was tripled (1 volt to 3 volts), the current increased ten times (5mA to 50 mA). At the same time the forward-bias voltage increased, the resistance decreased from 200 ohms to 60 ohms. In other words, when forward bias increases, the junction barrier gets smaller and this resistance to current flow decreases.

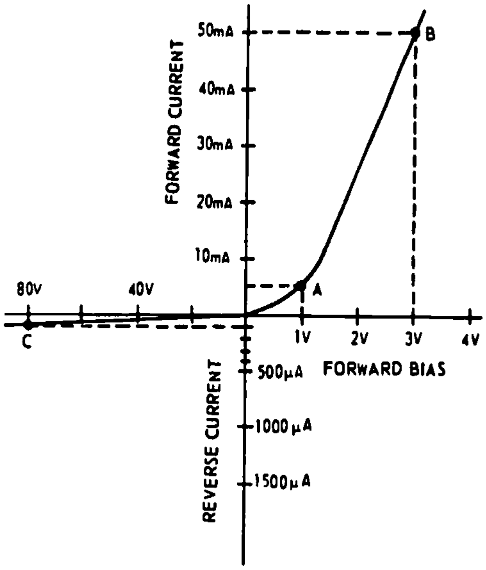


Fig 1-20. PN junction diode characteristics curve.

On the other hand, the diode conducts very little when reverse biased. Notice at point C the reverse bias voltage is 80 volts and the current is only 100 microamperes. This results in 800 k ohms of resistance, which is considerably larger than the resistance of the junction with forward bias. Because of these unusual features, the PN junction diode is often used to convert alternating current into direct current (rectification).

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is the purpose of a PN junction diode?

2. In reference to the schematic symbol for a diode, do electrons flow toward or away from the arrow?

3. What type of PN diode is formed by using a fine metal wire and a section of N-type semiconductor material?

4. Conduction in which type of semiconductor material is similar to conduction in a copper wire?

5. What is the name of the area in a PN junction that has a shortage of electrons and holes?

6. In order to reverse bias in a PN junction, what terminal of a battery is connected to the P terminal?

7. What type of bias opposes the PN junction barrier?

Work Unit 1-4. PN JUNCTION APPLICATION

DEFINE A LOAD AS REFERRED TO IN ELECTRONICS.

SPECIFY THE OUTPUT OF A HALF-WAVE RECTIFIER.

STATE THE TYPE OF BIAS WHICH MAKES A DIODE ACT AS A CLOSED SWITCH.

SPECIFY THE TYPE OF RECTIFIER THAT IS CONSTRUCTED BY SANDWICHING A SECTION OF SEMICONDUCTOR MATERIAL BETWEEN TWO METAL PLATES.

SPECIFY WHAT IS USED TO SHOW HOW DIODE PARAMETERS VARY OVER A FULL OPERATING RANGE.

DEFINE DIODE RATINGS.

Until now, we have mentioned only one application for the diode - rectification, but there are many more applications that we have not yet discussed. Variations in doping agents, semiconductor materials, and manufacturing techniques have made it possible to produce diodes that can be used in many different applications. Examples of these types of diodes are: signal diodes, rectifying diodes, zener diodes (voltage protection diodes for power supplies), varactors (amplifying and switching diodes), and many more. Only applications for two of the most commonly used diodes, the signal diode and rectifier diode, will be presented in this work unit. The other diodes will be explained later in this course.

One of the most important uses of a diode is rectification. The normal PN junction diode is well suited for this purpose as it conducts very heavily when forward biased (low-resistance direction) and only slightly when reverse biased (high-resistance direction). If we place this diode in series with a source of a.c. power, then the diode will be forward and reverse biased every cycle. Rectification is accomplished in this situation, when the current flows more easily in one direction than the other. The simplest rectifier circuit is a half-wave rectifier (fig 1-21) which consists of a diode, an a. c. power source, and a load resistor.

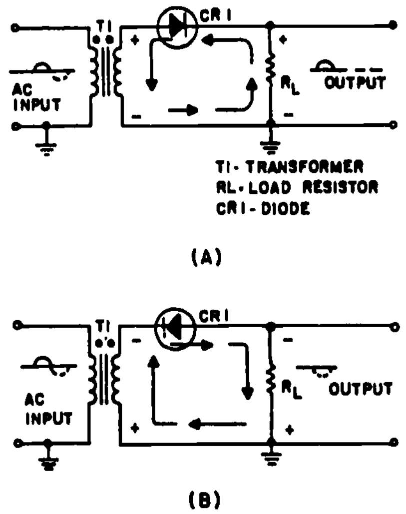


Fig 1-21. Simple half-wave rectifier.

The transformer (T1) in the figure provides the a.c. input to the circuit; the diode (CR1) provides the rectification; and the load resistor (R_L) serves two purposes: it limits the amount of current flow in the circuit to a safe level, and it also develops the output signal due to the current flow through it.

Before describing how this circuit operates, the definition of the word "load" as it applies to power supplies must be understood. Load is defined as any device that draws current. A device that draws little is considered a light load; whereas, a device that draws a large amount of current is a heavy load. Remember that when we speak of "load," we are speaking about the device that draws current from the power source. This device may be a simple resistor, or one or more complicated electronic circuits.

During the positive half-cycle of the input signal (solid line) in figure 1-21(A), the top of the transformer is positive with respect to ground. The dots on the transformer indicate points of the same polarity. With this condition, the diode is forward biased, the depletion region is narrow, the resistance of the diode is low, and current flows through the circuit in the direction of the solid lines. When this current flows through the load resistor, it develops a negative to positive voltage drop across it which appears as a positive voltage at the output terminal.

When the a.c. input goes in a negative direction, the top of the transformer becomes negative and the diode becomes reverse biased. With reverse bias applied to the diode, the depletion region increases, the resistance of the diode is high, and minimum current flows through the diode. For all practical purposes, there is no output developed across the load resistor during the negative alternation of the input signal as indicated by the broken lines in the figure. Although only one cycle of input is shown, it should be realized that the action described above continually repeats itself, as long as there is an input. Therefore, since only the positive half-cycles appear at the output, then this circuit converted the a.c. input into a positive pulsating d.c. voltage. The frequency of the output voltage is equal to the frequency of the applied a.c. signal since there is one pulse out for each cycle of the a.c. input. For example, if the input frequency is 60 hertz (60 cycles per second), the output frequency is 60 pulses per second (pps).

However, if the diode is reversed as shown in view (B) of figure 1-21, a negative output voltage would be obtained. This is because the current would be flowing from the top of R_L toward the bottom, making the output at the top of R_L negative in respect to the bottom or ground. Because current flows in this circuit only during half of the input cycle, it is called a half-wave rectifier.

The semiconductor diode shown in the figure can be replaced by a metallic rectifier and still achieve the same results. The metallic rectifier, sometimes referred to as a dry-disc rectifier, is a metal-to-semiconductor, large-area contact device. Its construction is distinctive; a semiconductor is sandwiched between two metal plates, or electrodes, as shown in figure 1-22. Note in the figure that a barrier, with a resistance many times greater than that of the semiconductor material, is constructed on one of the metal electrodes. The contact having the barrier is a rectifying contact; the other contact is nonrectifying. Metallic rectifiers act just like the diodes previously discussed in that they permit current to flow more readily in one direction than the other. However, the metallic rectifier is fairly large compared to the crystal diode as can be seen in figure 1-23. The reason for this is: metallic rectifier units are stacked (to prevent inverse voltage breakdown), have large area plates (to handle high currents), and usually have cooling fins (to prevent overheating).

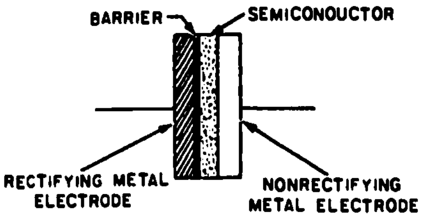


Fig 1-22. A metallic rectifier.

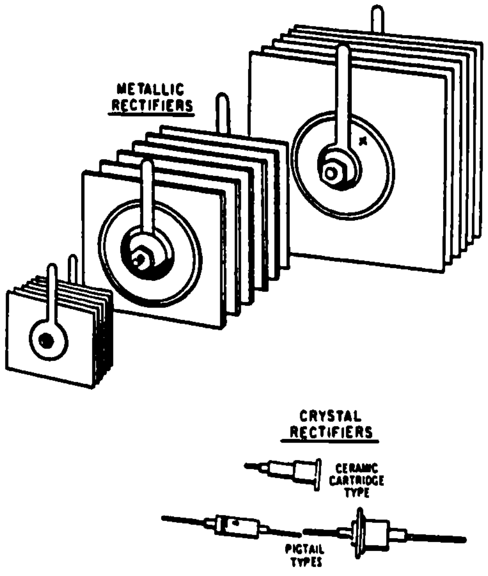


Fig 1-23. Different types of crystal and metallic rectifiers.

There are many known metal-semiconductor combinations that can be used for contact rectification. Copper oxide and selenium devices are by far the most popular. Copper oxide and selenium are frequently used over other types of metallic rectifiers because they have a large forward current per unit contact area, low forward voltage drop, good stability, and a lower aging rate. In practical applications, the selenium rectifier is used where a relatively large amount of power is required. On the other hand, copper-oxide rectifiers are generally used in small-current or for delivering direct current to circuits requiring not more than 10 amperes.

Since metallic rectifiers are affected by temperature, atmospheric conditions, and aging (in the case of copper oxide and selenium), they are being replaced by the improved silicon crystal rectifier. The silicon rectifier replaces the bulky selenium rectifier as to current and voltage rating, and can operate at higher ambient (surrounding) temperatures.

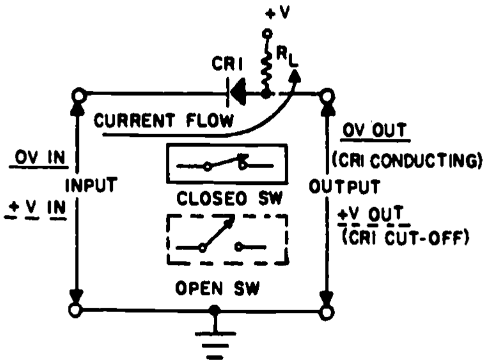


Fig 1-24. Basic diode switch.

In addition to their use as simple rectifiers, diodes are also used in circuits that mix signals together (mixers), detect the presence of a signal (detector), and act as a switch "to open or close a circuit." Diodes used in these applications are commonly referred to as "signal diodes." The simplest application of a signal diode is the basic diode switch shown in figure 1-24.

When the input to this circuit is at zero potential, the diode is forward biased because of the zero potential on the cathode and the positive voltage on the anode. In this condition, the diode conducts and acts as a straight piece of wire because of its very low forward resistance. In effect, the input is directly coupled to the output resulting in zero volts across the output terminals. Therefore, the diode acts as a closed switch when its anode is positive with respect to its cathode.

If we apply a positive input voltage (equal to or greater than the positive voltage supplied to the anode) to the diode's cathode, the diode will be reverse biased. In this situation, the diode is cut off and acts as an open switch between the input and output terminals. Consequently, with no current flow in the circuit, the positive voltage on the diode's anode will be felt at the output terminal. Therefore, the diode acts as an open switch when it is reverse biased.

Semiconductor diodes have properties that enable them to perform many different electronic functions. To do their jobs, engineers and technicians must be supplied with data on these different types of diodes. The information presented for this purpose is called DIODE CHARACTERISTICS. These characteristics are supplied by manufacturers either in their manuals or on specification sheets (data sheets). Because of the scores of manufacturers and numerous diode types, it is not practical to put before you a specification sheet and call it typical. Aside from the differences between manufacturers, a single manufacturer may even supply specification sheets that differ both in format and content. Despite these differences, certain performance and design information is normally required. It is this information that will be discussed in the next few paragraphs.

A standard specification sheet usually has a brief description of the diode. Included in this description is the type of diode, the major area of application, and any special features. Of particular interest is the specific application for which the diode is suited. The manufacturer also provides a drawing of the diode which gives dimensions, weight, and, if appropriate, any identification marks. In addition to the above data, the following information is also provided: a static operating table (giving spot values of parameters under fixed conditions), sometimes a characteristic curve similar to the one in figure 1-20 (showing how parameters vary over the full operating range), and diode ratings (which are the limiting values of operating conditions outside which could cause diode damage).

Manufacturers specify these various diode operating parameters and characteristics with "letter symbols" in accordance with fixed definitions. The following is a list, by letter symbol, of the major electrical characteristics for the rectifier and signal diodes.

RECTIFIER DIODES

D.C. BLOCKING VOLTAGE [V_R] - the maximum reverse d.c. voltage which will not cause breakdown.

AVERAGE FORWARD VOLTAGE DROP [$V_{F(AV)}$] - the average forward voltage drop across the rectifier given at a specified forward current and temperature.

AVERAGE RECTIFIER FORWARD CURRENT ($I_{F(AV)}$] - the average rectified forward current at a specified temperature, usually at 60 Hs with a resistive load.

AVERAGE REVERSE CURRENT [$I_{R(AV)}$] - the average reverse current at a specified temperature, usually at 60 Hz.

PEAK SURGE CURRENT [I_{surge}] - the peak current specified for a given number of cycles or portion of a cycle.

SIGNAL DIODES

PEAK REVERSE VOLTAGE [PRV] - the maximum reverse voltage which can be applied before reaching the breakdown point. (PRV also applies to the rectifier diode.)

REVERSE CURRENT [I_R] - the small value of direct current that flows when a semiconductor diode has reverse bias.

MAXIMUM FORWARD VOLTAGE DROP AT INDICATED FORWARD CURRENT [$V_F@I_F$] - the maximum forward voltage drop across the diode at the indicated forward current.

REVERSE RECOVERY TIME (t_{rr}) - the maximum time taken for the forward-bias diode to recover its reverse bias.

The electrical characteristics (by letter symbols) for the other types of diodes will appear later in this course.

The ratings of a diode (as stated earlier) are the limiting values of operating conditions which if exceeded could cause damage to a diode by either voltage breakdown or overheating. The PN junction diodes are generally rated for: MAXIMUM AVERAGE FORWARD CURRENT, PEAK RECURRENT FORWARD CURRENT, MAXIMUM SURGE CURRENT, and PEAK REVERSE VOLTAGE.

Maximum average forward current is usually given at a specified temperature, usually 25_0C, (77°F) and refers to the maximum amount of average current which can be permitted to flow in the forward direction. If this rating is exceeded, structure breakdown can occur.

Peak recurrent forward current is the maximum peak current which can be permitted to flow in the forward direction in the form of recurring pulses.

Maximum surge current is the maximum current permitted to flow in the forward direction in the form of non-recurring pulses. Current should not equal this value for more than a few milliseconds.

Peak reverse voltage(PRV) is one of the most important ratings. PRV indicates the maximum reverse-bias voltage which may be applied to a diode without causing junction breakdown.

All of the above ratings are subject to change with temperature variations. If, for example, the operating temperature is above that stated for the ratings, then the ratings must be decreased.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is a load?

____

2. What is the output of a half-wave rectifier?

3. What type of bias makes a diode act as a closed switch?

4. What type of rectifier is constructed by sandwiching a section of semiconductor material between two metal plates?

5. What is used to show how diode parameters vary over a full operating range?

6. What is meant by diode ratings?

Work Unit 1-5. DIODE IDENTIFICATION AND MAINTENANCE

IDENTIFY THE LETTER "N" AS USED IN THE SEMICONDUCTOR IDENTIFICATION SYSTEM.

NAME THE TYPES OF DIODES BY THE COLOR CODE SYSTEM.

SPECIFY THE GREATEST THREAT TO A DIODE.

STATE WHAT IS INDICATED BY TWO HIGH RESISTANCE MEASUREMENTS WHEN CHECKING A DIODE WITH AN OHMMETER.

There are many types of diodes varying in size from the size of a pinhead (used in subminiature circuitry) to large 250-ampere diodes (used in high power circuits). Because there are so many different types of diodes, then some system of identification is needed to distinguish one diode from another. This is accomplished with the semiconductor identification system shown in figure 1-25. This system is not only used for diodes but transistors and many other special semiconductor devices as well. As illustrated in this figure, the system uses numbers and letters to identify different types of semiconductor devices. The first number in the system indicates the number of junctions in the semiconductor device and is a number, one less than the number of active elements. Thus 1 designates a diode; 2 designates a transistor (which may be considered as made up of two diodes); and 3 designates a tetrode (a four-element transistor). The letter "N" following the first number indicates a semiconductor. The 2- or 3-digit number following the letter "N" is a serialized identification number. If needed, this number may contain a suffix letter after the last digit. For example, the suffix letter "M" may be used to describe matching pairs of separate semiconductor devices or the letter "R" may be used to indicate modified versions of the device which can be substituted for the basic numbered unit. For example, a semiconductor diode designated as type 1N345A signifies a two-element diode (1) of semiconductor material (N) that is an improved version (A) of type 345.

XNYYY

XN YYY

COMPONENT IDENTIFICATION NUMBER

X - NUMBER OF SEMICONDUCTOR JUNCTIONS
N - A SEMICONDUCTOR
YYY - IDENTIFICATION NUMBER (ORDER OR REGISTRATION NUMBER)
ALSO INCLUDES SUFFIX LETTER (IF APPLICABLE) TO INDICATE

1. MATCHING DEVICES
2. REVERSE POLARITY
3. MODIFICATION

EXAMPLE - IN345A (AN IMPROVED VERSION OF THE SEMICONDUCTOR DIODE TYPE 345)

Fig 1-25. Semiconductor identification code.

When working with these different types of diodes, it is also necessary to distinguish one end of the diode from the other (anode from cathode). For this reason, manufacturers generally code the cathode end of the diode with a "k", "+", "cath", "color dot or band", or by an unusual shape (raised edge or taper) as shown in figure 1-26. In some cases, standard color code bands are placed on the cathode end of the diode. This serves two purposes: (1) it identifies the cathode end of the diode, and (2) it also serves to identify the diode by number.

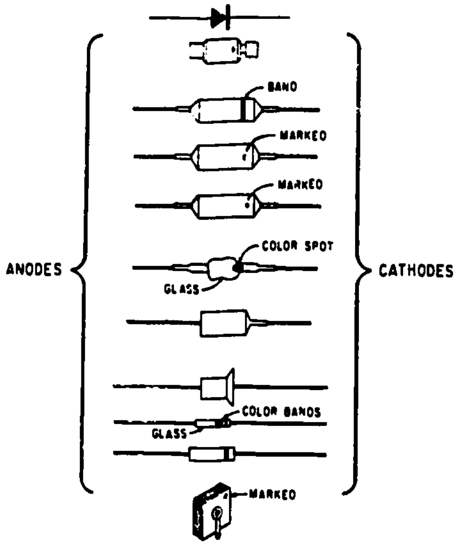
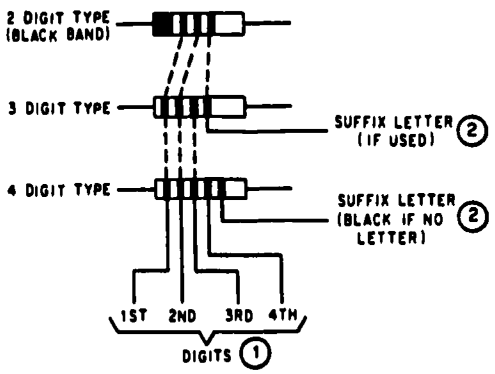


Fig 1-26. Semiconductor diode markings.



COLOR	① DIGIT	② DIODE SUFFIX LETTER
BLACK	0	-
BROWN	1	A
RED	2	B
ORANGE	3	C
YELLOW	4	D
GREEN	5	E
BLUE	6	F
VIOLET	7	G
GRAY	8	H
WHITE	9	J
SILVER	-	-
GOLD	-	-
NONE	-	-

Fig 1-27. Semiconductor diode color code system.

The standard diode color code system is shown in figure 1-27. Take, for example, a diode with brown, orange, and white bands at one terminal and figure out its identification number. With brown being a "1", orange a "3", and white "9", then the device would be identified as a type 139 semiconductor diode, or specifically 1N139.

Keep in mind, whether the diode is a small crystal type or a large power rectifier type, it is still represented schematically, as explained earlier, by the schematic symbol shown in figure 1-12.

Diodes are rugged and efficient. They are also expected to be relatively trouble free. Protective encapsulation processes and special coating techniques have even further increased their life expectancies. In theory, a diode should last indefinitely. However, if diodes are subjected to current overloads, their junctions will be damaged or even destroyed. In addition, the application of excessively high operating voltages can damage or destroy junctions through arc-over, or excessive reverse currents. One of the greatest dangers to the diode is heat. Heat causes more electron-hole pairs to be generated which, in turn, increases current flow. This increase in current generates more heat and the cycle repeats itself until the diode draws excessive current. This action is referred to as THERMAL RUNAWAY and eventually causes diode destruction. Extreme caution should be used when working with equipment containing diodes to ensure that these problems do not occur and cause irreparable diode damage.

The following is a list of some of the special safety precautions which should be observed when working with diodes.

Never remove or insert a diode into a circuit with voltage applied.

Never pry diodes to loosen them from their circuits.

Always be careful when soldering to ensure that excessive heat is not applied to the diode.

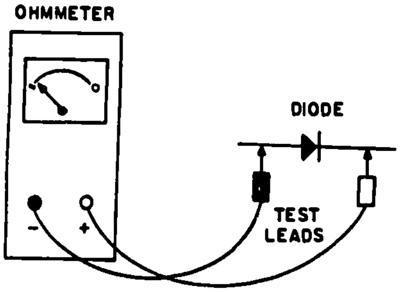
When testing a diode, ensure that the test voltage does not exceed the diode's maximum allowable voltage.

Never put your fingers across a signal diode because the static charge from your body could short it out.

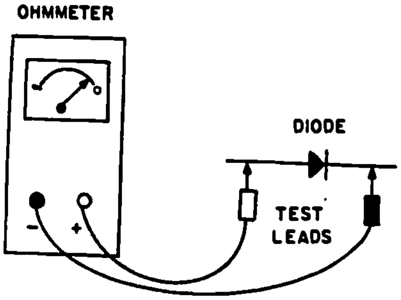
Always replace a diode with one of the same type or with a direct replacement.

Ensure a replacement diode is put into a circuit in the correct direction.

If a diode has been subjected to excessive voltage or temperature and is suspected of being defective, it can be checked in various ways. <u>The most convenient and quickest way of testing a diode is with an ohmmeter (figure 1-28).</u> <u>To make this check, simply disconnect one of the diode leads from the circuit wiring</u>, and make resistance measurements across the leads of the diode. The resistance measurements obtained depend upon the test-lead polarity of the ohmmeter; therefore, two measurements must be taken. The first measurement is taken with the test leads connected to either end of the diode and the second measurement is taken with the test leads reversed on the diode. The <u>larger resistance</u> value is assumed to be the <u>reverse (back) resistance</u> of the diode, and the <u>smaller resistance</u> (front) value is assumed <u>to be the forward resistance</u>. Measurements can be <u>made for comparison purposes</u> using another <u>identical-type diode</u>, known to be good, as a <u>standard</u>. <u>Two high-value</u> resistance measurements indicate that the diode is open or has a high forward resistance; two low-value resistance measurements indicate that the diode is shorted or has a low reverse resistance. A normal set of measurements will show a <u>high resistance</u> in the <u>reverse direction</u> and a <u>low resistance</u> in the <u>forward direction</u>. The <u>diode's efficiency</u> is <u>determined by how</u> the <u>forward resistance</u> is as compared with the reverse resistance; that is, it is desirable to have as great a ratio (often known as the front to back ratio or the back to front ratio) as possible between the reverse and forward resistance measurements. However, as a rule of thumb, a small signal diode will have a ratio of several hundred to one, while a power rectifier can operate satisfactorily with a ratio of 10 to 1.



REVERSE CONDITION -
HIGH RESISTANCE MEASUREMENT



FORWARD CONDITION -
LOW RESISTANCE MEASUREMENT

Fig. 1-28. Checking a diode with an ohmmeter.

One thing you should keep in mind about the ohmmeter check _ it is <u>not</u> conclusive. It is still possible for a diode to check good under this test, but breaks down when replaced in the circuit. The problem is that the meter used to check the diode uses a lower voltage than that which the diode usually operates at in the circuit. Another important point to remember is that a diode should not be condemned because two ohmmeters give different readings on the diode. This occurs due to the different internal resistance of the ohmmeters and the different states of charge on the ohmmeter batteries. Because each ohmmeter sends a different current through the diode, the two resistance value readings on the meters will not be the same.

Another way of checking a diode is with the substitution method. In this method, a good diode is substituted for a questionable diode. This technique should be used only after you have made voltage and resistance measurements to make certain that there is no circuit defect that might damage the substitution diode. If more than one defective diode is present in the equipment section where trouble has been localized, this method becomes cumbersome, since several diodes may have to be replaced before the trouble is corrected. To determine which stages failed and which diodes are not defective, all of the removed diodes must be tested. This can be accomplished by observing whether the equipment operates correctly as each of the removed diodes is reinserted into the equipment.

In conclusion, the only valid check of a diode is a dynamic electrical test which determines the diode's forward current (resistance) and reverse current (resistance) parameters. This test can be accomplished using various crystal diode test sets.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What does the letter "N" indicate in the semiconductor identification system?

____

2. What type of diode has orange, blue, and gray bands?

____

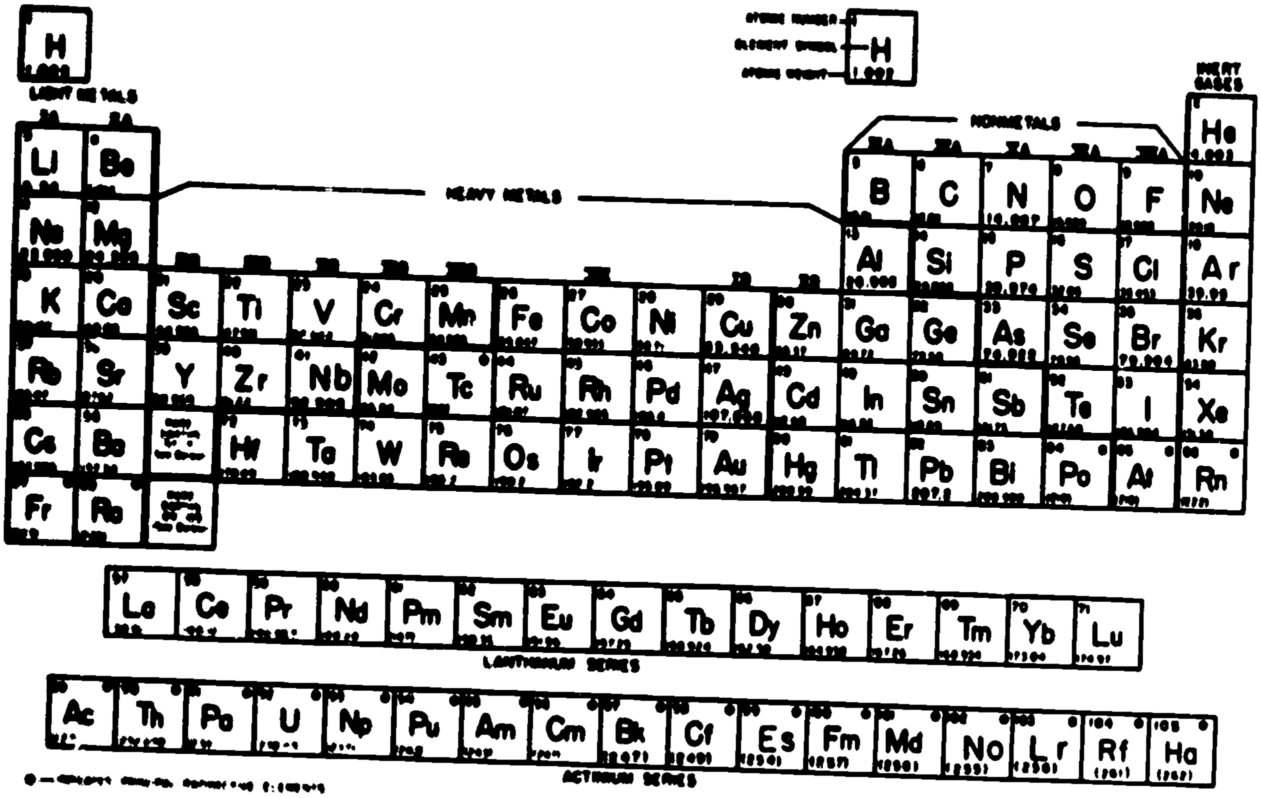
3. What is the greatest threat to a diode?

____

4. When checking a diode with an ohmmeter, what is indicated by two high resistance measurements?

____

PERIODIC TABLE OF THE ELEMENTS



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	[illegible]	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	[illegible]															

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu		
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	Rf	Ha

SEE NEXT PAGE FOR INTERPRETATION OF SYMBOLS

Table 1-1. Periodic Table of the Elements - Continued

Symbol	Name	Atomic Number	Atomic Weight
Ac	Actinium	89	1(227)
Ag	Silver	47	107.868
Al	Aluminum	13	26.982
Am	Americium	95	(243)
Ar	Argon	18	39.95
As	Arsenic	33	74.922
At	Astatine	85	(210)
Au	Gold	79	196.967
B	Boron	5	10.81
Ba	Barium	56	137.34
Be	Beryllium	4	9.012
Bi	Bismuth	83	208.980
Bk	Berkelium	97	(247)
Br	Bromine	35	79.904
C	Carbon	6	12.011
Ca	Calcium	20	40.08
Cd	Cadmium	48	112.40
Ce	Cerium	58	140.12
Cf	Californium	98	(249)
Cl	Chlorine	17	35.453
Cm	Curium	96	(247)
Co	Cobalt	27	58.933
Cr	Chromium	24	51.996
Cs	Cesium	55	132.905
Cu	Copper	29	63.546
Dy	Dysprosium	66	162.50
Es	Einsteinium	99	(254)
Er	Erbium	68	167.26
Eu	Europium	63	151.96
F	Fluorine	9	18.998
Fe	Iron	26	55.847
Fm	Fermium	100	(257)
Fr	Francium	87	(223)
Ga	Gallium	31	69.72
Gd	Gadolinium	64	157.25
Ge	Germanium	32	72.59
H	Hydrogen	1	1.008
Ha	Hahnium	105	(262)
He	Helium	2	4.003
Hf	Hafnium	72	178.49
Hg	Mercury	80	200.59
Ho	Holmium	67	164.930
I	Iodine	53	126.904
In	Indium	49	114.82
Ir	Iridium	77	192.2
K	Potassium	19	39.102
Kr	Krypton	36	83.80
La	Lanthanum	57	138.91
Li	Lithium	3	6.94
Lr	Lawrencium	103	(256)
Lu	Lutetium	71	174.97
Md	Mendelevium	101	(258)
Mg	Magnesium	12	24.305

Table 1-1. Periodic Table of the Elements - Continued

Symbol	Name	Atomic Number	Atomic Weight
Mn	Manganese	25	54.938
Mo	Molybdenum	42	95.94
N	Nitrogen	7	14.007
Na	Sodium	11	22.990
Nb	Niobium	41	92.906
Nd	Neodymium	60	144.24
Ne	Neon	10	20.18
Ni	Nickel	28	58.71
No	Nobelium	102	(255)
Np	Neptunium	93	(237)
O	Oxygen	8	15.999
Os	Osmium	76	190.2
P	Phosphorus	15	30.974
Pa	Protactinium	91	(231)
Pb	Lead	82	207.2
Pd	Palladium	46	106.4
Pm	Promethium	61	(147)
Po	Polonium	84	(210)
Pr	Praseodymium	59	140.907
Pt	Platinum	78	195.09
Pu	Plutonium	94	(242)
Ra	Radium	88	(226)
Rb	Rubidium	37	85.47
Re	Rhenium	75	186.2
Rf	Rutherfordium	104	(261)
Rh	Rhodium	45	102.905
Rn	Radon	86	(222)
Ru	Ruthenium	44	101.07
S	Sulfur	16	32.06
Sb	Antimony	51	121.75
Sc	Scandium	21	44.956
Se	Selenium	34	78.96
Si	Silicon	14	28.086
Sm	Samarium	62	150.35
Sn	Tin	50	118.69
Sr	Strontium	38	87.62
Ta	Tantalum	73	180.948
Tb	Terbium	65	158.924
Tc	Technetium	43	(99)
Te	Tellurium	52	127.60
Th	Thorium	90	232.038
Ti	Titanium	22	47.90
Tl	Thallium	81	204.37
Tm	Thulium	69	168.934
U	Uranium	92	238.03
V	Vanadium	23	50.942
W	Tungsten	74	183.85
Xe	Xenon	54	131.30
Y	Yttrium	39	88.905
Yb	Ytterbium	70	173.04
Zn	Zinc	30	65.37
Zr	Zirconium	40	91.22

SUMMARY REVIEW

Now that you have completed this study unit, you have learned the history of semiconductors, their theory of operation, their purpose, and application, and how to identify and maintain them. If you had any problems it is highly recommended that you go over this study unit prior to proceeding.

Answer to Study Unit #1 Exercises

Work Unit 1-1.

1. An electrical device which operates by the movement of electrons within a solid piece of semiconductor material.
2. It is the decrease in a semiconductor resistance as temperature rises.
3. a. Space systems
   b. Computers
   c. Data processing equipment
4. The electron tube requires filament or heater voltage; whereas, the semiconductor device does not. Consequently, no power input is spent by the semiconductor for conduction.

Work Unit 1-2.

1. Anything that occupies space and has weight
2. a. Solid
   b. Liquid
   c. Gas
3. The valence shell
4. The width of the forbidden band
5. Quanta
6. Covalent bonding
7. a. Electron flow
   b. Hole flow
8. Intrinsic
9. P-type crystal
10. Electrons

Work Unit 1-3.

1. To convert alternating current into direct current
2. Toward the arrow
3. Point contact
4. N-type material
5. Depletion region
6. Negative terminal
7. Forward

Work Unit 1-4.

1. Any device that draws current
2. A pulsating d.c. voltage
3. Forward bias
4. Metallic rectifier
5. A characteristic curve
6. They are the limiting values of operating conditions outside which operations could cause diode damage.

Work Unit 1-5.

1. A semiconductor
2. 1N368
3. Heat
4. That the diode is open or has a high forward resistance

STUDY UNIT 2

TRANSISTORS

STUDY UNIT OBJECTIVE: WITHOUT THE AID OF REFERENCES, YOU WILL IDENTIFY THE VARIOUS TYPES OF TRANSISTORS AND THEIR CONSTRUCTION, ALPHANUMERICAL DESIGNATIONS, SYMBOLS, AND OPERATION. YOU WILL IDENTIFY THE PRECAUTIONS THAT MUST BE TAKEN WHEN WORKING WITH TRANSISTORS. YOU WILL IDENTIFY INTEGRATED CIRCUITS, THEIR CONSTRUCTION, AND ADVANTAGES THEY AFFORD THEIR USERS. LASTLY, YOU WILL IDENTIFY THE FUNCTION OF MODULAR CIRCUITRY.

The discovery of the first transistor in 1948 by a team of physicists at the Bell Telephone Laboratories sparked an interest in solid-state research that spread rapidly. The transistor, which began as a simple laboratory oddity, was rapidly developed into a semiconductor device of major importance. The transistor demonstrated for the first time in history that amplification in solids was possible. Prior to the transistor, amplification was achieved only with electron tubes. Transistors now perform numerous electronic tasks with new and improved transistor designs being continually put on the market. In many cases, transistors are more desirable than tubes because they are small, rugged, require no filament power, and operate at low voltages with comparatively high efficiency. The development of a family of transistors has even made possible the miniaturization of electronic circuits. Figure 2-1 shows a sample of the many different types of transistors you may encounter when working with electronic equipment.

Transistors have infiltrated virually every area of science and industry, from the family car to satellites. The military depends heavily on transistors. The ever increasing uses for transistors have created an urgent need for sound and basic information regarding their operation.

From your study of the PN-junction diode in the preceding topic, you now have the basic knowledge to grasp the principles of transistor operation. In this topic you will first become acquainted with the basic types of transistors, their construction, and their theory of operation. You will also find out just how and why transistors amplify. Once this basic information is understood, transistor terminology, capabilities, limitations, and identification will be discussed. Last, we will talk about transistor maintenance, integrated circuits, boards, and modular circuitry.

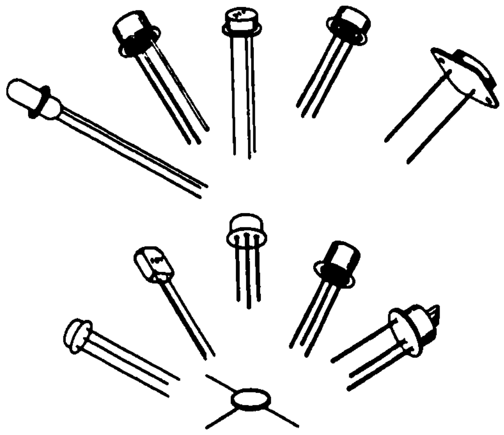


Fig 2-1. An assortment of different types of transistors.

Work Unit 2-1. TRANSISTOR FUNDAMENTALS

SPECIFY THE SEMICONDUCTOR DEVICE THAT HAS THREE OR MORE ELEMENTS.

STATE THE ELECTRONIC FUNCTION WHICH MAKES THE TRANSISTOR SPECIAL.

NAME THE VERY FIRST TRANSISTOR.

SPECIFY ONE OF THE MOST IMPORTANT PARTS OF ANY TRANSISTOR MANUFACTURING PROCESS.

The first solid-state device discussed was the two-element semiconductor diode. The next device on our list is even more unique. It not only has one more element than the diode but it can amplify as well. Semiconductor devices that have three or more elements are called TRANSISTORS. The term transistor was derived from the words TRANSfer and resISTOR. This term was adopted because it best describes the operation of the transistor--the transfer of an input signal current from a low-resistance circuit to a high-resistance circuit. Basically, the transistor is a solid-state device that amplifies by controlling the flow of current carriers through its semiconductor materials.

There are many different types of transistors, but their basic theory of operation is all the same. As a matter of fact, the theory we will be using to explain the operation of a transistor is the same theory used earlier with the PN-junction diode except that now two such junctions are required to form the three elements of a transistor. The three elements of the two junction transistor are: (1) the EMITTER, which gives off, or "emits," current carriers (electrons or holes); (2) the BASE, which controls the flow of current carrier; and (3) the COLLECTOR, which collects the current carriers.

Transistors are classified as either NPN or PNP according to the arrangement of their N and P materials. Their basic construction and chemical treatment is implied by their names, "NPN" or "PNP." That is, an NPN transistor is formed by introducing a thin region of P-type material between two regions of N-type material. On the other hand, a PNP transistor is formed by introducing a thin region of N-type material between two regions of P-type material. Transistors constructed in this manner have two PN junctions, as shown in figure 2-2. One PN junction is between the emitter and the base; the other PN junction is between the collector and the base. The two junctions share one section of semiconductor material so that the transistor actually consists of three elements.

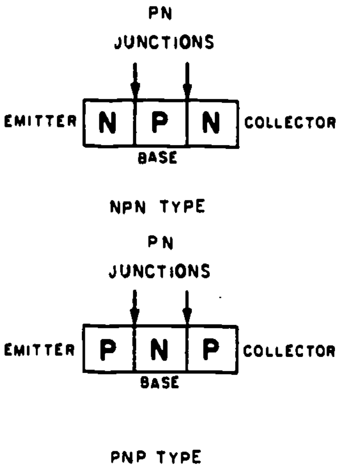


Fig 2-2. Transistor block diagrams.

Since the majority and minority current carriers are different for N and P materials, it stands to reason that the internal operation of the NPN and PNP transistors will also be different. The theory of operation of the NPN and PNP transistors will be discussed separately in the next few paragraphs. Any additional information about the PN junction will be given as the theory of transistor operation is developed.

In order to prepare you for the forthcoming information, the two basic types of transistors along with their circuit symbols are shown in figure 2-4. It should be noted that the two symbols are different. The horizontal line represents the base, the angular line with the arrow on it represents the collector. The direction of the arrow (Fig 2-3) on the emitter distinguishes the NPN from the PNP transistor. If the arrow points in, (Points IN) the transistor is a PNP. On the other hand, if the arrow points out, the transistor is an NPN (Not Pointing iN).

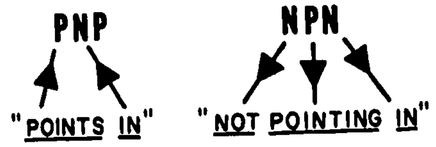


Fig 2-3. Distinguishing the NPN from the PNP transistor.

Another point you should keep in mind is that the arrow always points into the direction of hole flow, or from the P to N sections, no matter whether the P section is the emitter or base. On the other hand, electron flow is always toward or against the arrow, just like in the junction diode. The very first transistors were known as point-contact transistors. Their construction is similar to the construction of the point-contact diode covered in study unit 1. The difference, of course, is that the point-contact transistor has two P or N regions formed instead of one. Each of the two regions constitutes an electrode (element) of the transistor. One is named the emitter and the other is named the collector, as shown in figure 2-5, view (A).

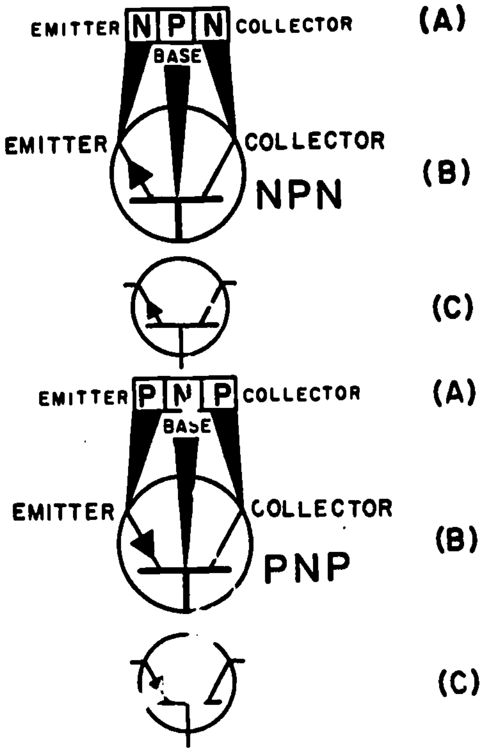


Fig 2-4. Transistor representations.

Point-contact transistors are now practically obsolete. They have been replaced by junction transistors, which are superior to point-contact transistors in nearly all respects. The junction transistor generates less noise, handles more power, provides higher current and voltage gains, and can be mass-produced more cheaply than the point-contact transistor. Junction transistors are manufactured in much the same manner as the PN-junction diode discussed earlier. However, when the PNP or NPN material is grown (view B), the impurity mixing process must be reversed twice in order to obtain the two junctions required in a transistor. Likewise, when the alloy-junction (view C) or the diffused-junction (view D) process is used, two junctions must also be created within the crystal.

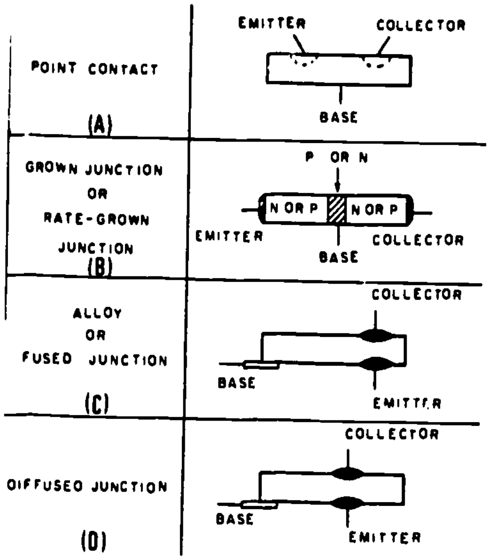


Fig 2-5. Transistor constructions.

Although there are numerous ways to manufacture transistors, one of the most important parts of any manufacturing process is quality control. Without good quality control, many transistors would prove unreliable because the construction and processing of a transistor govern its thermal ratings, stability, and electrical characteristics. Even though there are many variations in the transistor manufacturing processes, certain structural techniques, which yield good reliability and long life, are common to all processes: (1) Wire leads are connected to each semiconductor electrode; (2) the crystal is specially mounted to protect it against mechanical damage; and (3) the unit is sealed to prevent harmful contamination of the crystal.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. Specify the name given to the semiconductor device that has three or more elements.

2. State the electronic function which makes the transistor special?

3. Name the very first transistor.

4. Specify one of the most important parts of any transistor manufacturing process.

Work Unit 2-2. TRANSISTOR THEORY

SPECIFY WHAT POLARITY VOLTAGE IS APPLIED TO THE COLLECTOR, AND ITS RELATIONSHIP TO THE BASE VOLTAGE TO PROPERLY BIAS AN NPN TRANSISTOR.

STATE THE PERCENTAGE OF CURRENT IN AN NPN TRANSISTOR THAT REACHES THE COLLECTOR.

STATE THE RELATIONSHIP BETWEEN THE POLARITY OF THE VOLTAGE APPLIED TO THE PNP TRANSISTOR AND THE RELATIONSHIP APPLIED TO THE NPN TRANSISTOR.

NAME THE TWO CURRENT LOOPS IN A TRANSISTOR.

NAME THE LETTER DESIGNATION FOR BASE CURRENT.

You should recall from an earlier discusion that a forward-biased PN junction is comparable to a low-resistance circuit element because it passes a high current for a given voltage. In turn, a reverse-biased PN junction is comparable to a high-resistance circuit element. By using the Ohm's law formula for power ($P = I^2R$) and assuming current is held constant, you can conclude that the power developed across a high resistance is greater than that developed across a low resistance. Thus, if a crystal was to contain two PN junctions (one forward-biased and the other reverse-biased), a low-power signal could be injected into the forward-biased junction and produce a high-power signal at the reverse-biased junction. In this manner, a power gain would be obtained across the crystal. This concept, which is merely an extension of the material covered in study unit 1, is the basic theory behind how the transistor amplifies. With this information fresh in your mind, let's proceed directly to the NPN transistor.

Just as in the case of the PN junction diode, the N material comprising the two end sections of the NPN transistor contains a number of free electrons, while the center P section contains an excess number of holes. The action at each junction between these sections is the same as that previously described for the diode. That is, depletion regions develop and the junction barrier appears. In order to use the transistor as an amplifier, each of these junctions must be modified by some external bias voltage. For the transistor to function in this capacity, the first PN junction (emitter-base junction) is biased in the forward, or low-resistance, direction. At that same time the second PN junction (base-collector junction) is biased in the reverse, or high-resistance, direction. A simple way to remember how to properly bias a transistor is to observe the NPN or PNP elements that make up the transistor. The letters of these elements indicate what polarity voltage to use for correct bias. For instance, notice the NPN transistor below: shown in figure 2-6.

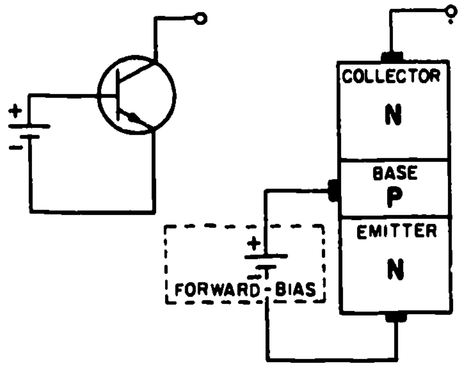


Fig 2-6. NPN Transistor.

a. The emitter, which is the first letter in the NPN sequence, is connected to the negative side of the battery while the base, which is the second letter (NPN), is connected to the positive side.

b. However, since the second PN junction is required to be reverse biased for proper transistor operation, then the collector must be connected to an opposite polarity voltage (positive) than that indicated by its letter designation (NPN). The voltage on the collector must also be more positive than the base, shown in figure 2-7.

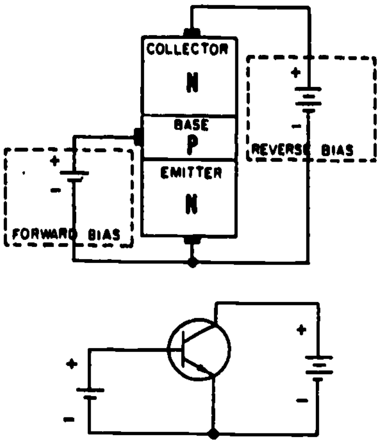


Fig 2-7 Properly biased NPN transistor.

We now have a properly biased NPN transistor.

In summary, the base of the NPN transistor must be positive with respect to the emitter, and the collector must be more positive than the base.

NPN FORWARD-BIASED JUNCTION (Fig 2-8).--An important point to bring out at this time, which was not mentioned during the explanation of the diode, is the fact that the N material on one side of the forward-biased junction is more heavily doped than the P material. This results in more current being carried across the junction by the majority carrier electrons from the N material than the majority carrier holes from the P material. Therefore, conduction through the forward-biased junction, as shown in figure 2-8, is mainly by majority carrier electrons from the N material (emitter).

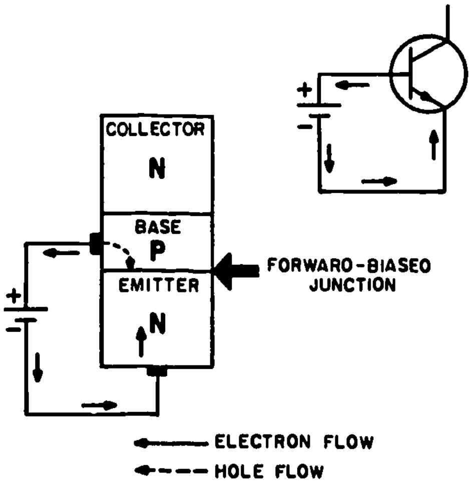


Fig 2-8. The forward-biased junction in an NPN transistor.

With the emitter-to-base junction in the figure biased in the forward direction, electrons leave the negative terminal of the battery and enter the N material (emitter). Since electrons are majority current carriers in the N material, they pass easily through the emitter, cross over the junction, and combine with holes in the P material (base). For each electron that fills a hole in the P material, another electron will leave the P material (creating a new hole) and enter the positive terminal of the battery.

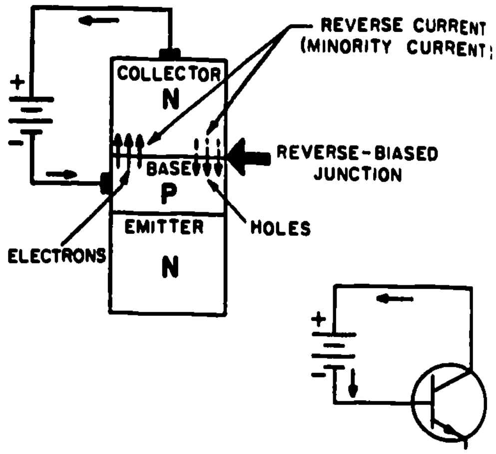


Fig 2-9. The reverse-biased junction in an NPN transistor.

NPN REVERSE-BIASED JUNCTION. (Fig 2-9)--The second PN junction (base-to-collector), or reverse-biased junction as it is called (figure 2-9), blocks the majority current carriers from crossing the junction. However, there is a very small current, mentioned earlier, that does pass through this junction. This current is called minority current, or reverse current. As you recall, this current was produced by the electron-hole pairs. The minority carriers for the reverse-biased PN junction are the electrons in the P material and the holes in the N material. These minority carriers actually conduct the current for the reverse-biased junction when electrons from the P material enter the N material, and the holes from the N material enter the P material. However, the minority current electrons (as you will see later) play the most important part in the operation of the NPN transistor.

At this point you may wonder why the second PN junction (base-to-collector) is not forward biased like the first PN junction (emitter-to-base). If both junctions were forward biased, the electrons would have a tendency to flow from each end section of the N P N transistor (emitter and collector) to the center P section (base). In essence, we would have two junction diodes possessing a common base, thus eliminating any amplification and defeating the purpose of the transistor. A word of caution is in order at this time. If you should mistakenly bias the second PN junction in the forward direction, the excessive current could develop enough heat to destroy the junctions, making the transistor useless. Therefore, be sure your bias voltage polarities are correct before making any electrical connections.

NPN JUNCTION INTERACTION (Fig 2-10)-We are now ready to see what happens when we place the two junctions of the NPN transistor in operation at the same time. For a better understanding of just how the two junctions work together, refer to figure 2-10 during the discussion.

The bias batteries in this figure have been labeled V_{CC} for the collector voltage supply, and V_{BB} for the base voltage supply. Also notice the base supply battery is quite small, as indicated by the number of cells in the battery, usually 1 volt or less. However, the collector supply is generally much higher than the base supply (normally around 6 volts). As you will see later, this difference in supply voltages is necessary in order to have current flow from the emitter to the collector.

As stated earlier, the current flow in the external circuit is always due to the movement of free electrons. Therefore, electrons flow from the negative terminals of the supply batteries to the N-type emitter. This combined movement of electrons is known as emitter current (I_E). Since electrons are the majority carriers in the N material, they will move through the N material emitter to the emitter-base junction. With this junction forward biased, electrons continue on into the base region. Once the electrons are in the base, which is a P-type material, they now become minority carriers. Some of the electrons that move into the base recombine with available holes. For each electron that recombines, another electron moves out through the base lead as base current I_B (creating a new hole for eventual combination) and returns to the base supply battery V_{BB}. The electrons that are combined are lost as far as the collector is concerned. Therefore, in order to make the transistor more efficient, the base region is made very thin and lightly doped. This reduces the opportunity for an electron to recombine with a hole and be lost. Thus, most of the

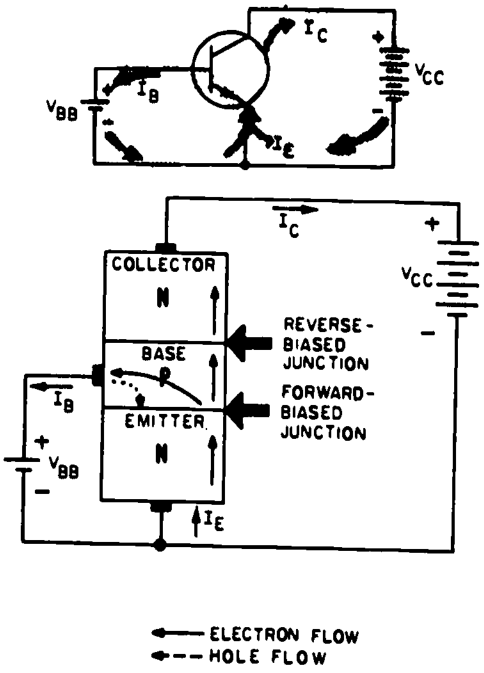


Fig 2-10. NPN transistor operation.

electrons that move into the base region come under the influence of the large collector reverse bias. This bias acts as forward bias for the minority carriers (electrons) in the base and, as such, accelerates them through the base-collector junction and on into the collector region. Since the collector is made of an N-type material, the electrons that reach the collector again become majority current carriers. Once in the collector, the electrons move easily through the N material and return to the positive terminal of the collector supply battery V_{CC} as collector current (I_C). To further improve on the efficiency of the transistor, the collector is made physically larger than the base for two reasons: (1) to increase the chance of collecting carriers that diffuse to the side as well as directly across the base region, and (2) to enable the collector to handle more heat without damage. In summary, total current flow in the NPN transistor is through the emitter lead. Therefore, in terms of percentage, I_E is 100 percent. On the other hand, since the base is very thin and lightly doped, a smaller percentage of the total current (emitter current) will flow in the base circuit than in the collector circuit. Usually no more than 2 to 5 percent of the total current is base current (I_B) while the remaining 95 to 98 percent is collector current (I_C). A very basic relationship exists between these two currents:

$$I_E = I_B + I_C$$

In simple terms this means that the emitter current is separated into base and collector current. Since the amount of current leaving the emitter is solely a function of the emitter-base bias, and because the collector receives most of this current, a small change in emitter-base bias will have a far greater effect on the magnitude of collector current than it will have on base current. In conclusion, the relatively small emitter-base bias controls the relatively large emitter-to-collector current.

The PNP transistor works essentially the same as the NPN transistor. However, since the emitter, base, and collector in the PNP transistor are made of materials that are different from those used in the NPN transistor, different current carries flow in the PNP unit. The majority current carriers in the PNP transistor are holes. This is in contrast to the NPN transistor where the majority current carriers are electrons. In order to support this different type of current (hole flow), the bias batteries are reversed for the PNP transistor. A typical bias setup for the PNP transistor is shown in figure 2-11. Notice that the procedure used earlier to properly bias the NPN transistor also applies here to the PNP transistor. The first letter (P) in the PNP sequence indicates the polarity of the voltage required for the emitter (positive), and the second letter (N) indicates the polarity of the base voltage (negative). Since the base-collector junction is always reverse biased, then the opposite polarity voltage (negative) must be used for the collector. Thus, the base of the PNP transistor must be negative with respect to the emitter, and the collector must be more negative than the base. Remember, just as in the case of the NPN transistor, this difference in supply voltage is necessary in order to have current flow (hole flow in the case of the PNP transistor) from the emitter to the collector. Although hole flow is the predominant type of current flow in the PNP transistor, hole flow only takes place within the transistor itself, while electrons flow in the external circuit. However, it is the internal hole flow that leads to electron flow in the external wires connected to the transistor.

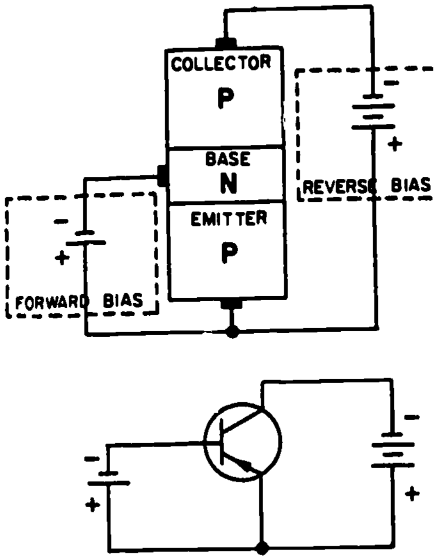


Fig 2-11. A properly biased PNP transistor.

PNP FORWARD-BIASES JUNCTION (Fig 2-12)-Now let's consider what happens when the emitter-base junction in figure 2-12 is forward biased. With the bias setup shown, the positive terminal of the battery repels the emitter holes toward the base, while the negative terminal drives the base electrons toward the emitter. When an emitter hole and a base electron meet, they combine. For each electron that combines with a hole, another electron leaves the negative terminal of the battery, and enters the base. At the same time, an electron leaves the emitter, creating a new hole, and enters the positive terminal of the battery. This movement of electrons into the base and out of the emitter constitutes base current flow (I_B), and the path these electrons take is referred to as the emitter-base circuit.

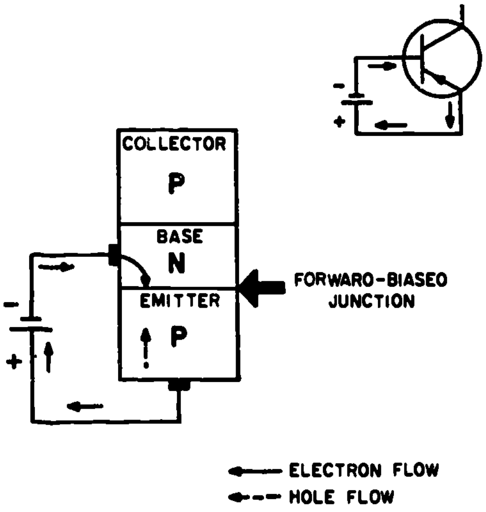


Fig 2-12. The forward-biased junction in a PNP transistor.

PNP REVERSE-BIASED JUNCTION (Fig 2-13)--In the reverse-biased junction (figure 2-13), the negative voltage on the collector and the positive voltage on the base block the majority current carriers from crossing the junction. However, this same negative collector voltage acts as forward bias for the minority current holes in the base, which cross the junction and enter the collector. The minority current electrons in the collector also sense forward bias--the positive base voltage--and move into the base. The holes in the collector are filled by electrons that flow from the negative terminal of the battery. At the same time the electrons leave the negative terminal of the battery, other electrons in the base break their covalent bonds and enter the positive terminal of the battery. Although there is only minority current flow in the reverse-biased junction, it is still very small due to the limited number of minority current carriers.

PNP JUNCTION INTERACTION.--The interaction between the forward - and reverse-biased junctions in a PNP transistor is very similar to that in an NPN transistor, except that in the PNP transistor, the majority current carriers are holes. In the PNP transistor, shown in figure 2-14, the positive voltage on the emitter repels the holes toward the base. Once in the base, the holes combine with base electrons, but again, remember that the base region is made very thin to prevent the recombination of holes with electrons. Therefore, well over 90 percent of the holes that enter the base become attracted to the large negative collector voltage and pass right through the base. However, for each electron and hole that combine in the base region, another electron leaves the negative terminal of the base battery (V_{BB}) and enters the base as base current (I_B). At the same time an electron leaves the negative terminal of the battery, another electron leaves the emitter as I_E (creating a new hole) and enters the positive terminal of V_{BB}. Meanwhile, in the collector circuit, electrons from the collector battery (V_{CC}) enter the collector as I_C and combine with the excess holes from the base. For each hole that is neutralized in the collector by an electron, another electron leaves the emitter and starts its way back to the positive terminal of V_{CC}.

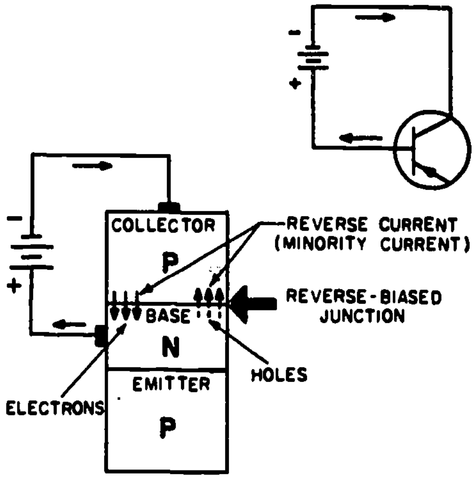


Fig 2-13. The reversed-biased junction in a PNP transistor.

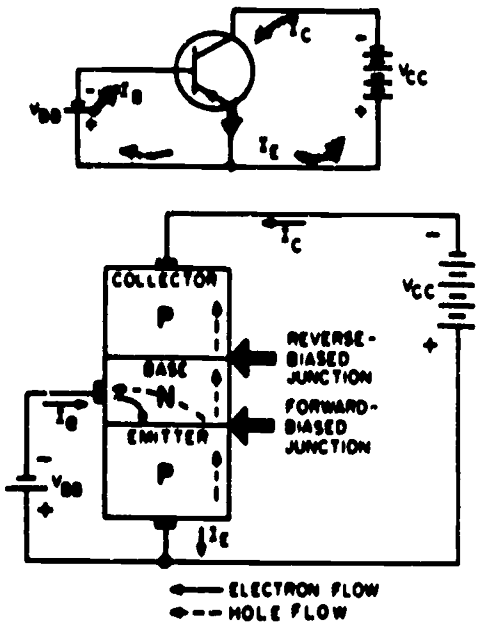


Fig 2-14. PNP transistor operation.

Although current flow in the external circuit of the PNP transistor is opposite in direction to that of the NPN transistor, the majority carriers always flow from the emitter to the collector. This flow of majority carriers also results in the formation of two individual current loops within each transistor. One loop is the base-current path and the other loop the collector-current path. The combination of the current in both of these loops (I_B + I_C) results in total transistor current (I_E). The most important thing to remember about the two different types of transistors is that the emitter-base voltage of the PNP transistor has the same controlling effect on collector current as that of the NPN transistor. In simple terms, increasing the forward-bias voltage of a transistor reduces the emitter-base junction barrier. This action allows more carriers to reach the collector causing an increase in current flow from the emitter to the collector and through the external circuit. Conversely, a decrease in the forward-bias voltage reduces collector current.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. To properly bias an NPN transistor, what polarity voltage is applied to the collector?

____

2. Referring to question #1, what is its relationship to the base voltage?

____

3. Why is conduction through the forward-biased junction of an NPN transistor primarily in one direction (namely from the emitter to base)?

____

____

4. In the NPN transistor, what section is made very thin compared with the other two sections?

____

5. What percentage of current in an NPN transistor reaches the collector?

____

6. What are the majority current carriers in a PNP transistor?

____

7. What is the relationship between the polarity of the voltage applied to the PNP transistor and that applied to the NPN transistor?

____

8. What is the letter designation for base current?

____

9. Name the two current loops in a transistor. ____________________

____

Work Unit 2-3. BASIC TRANSISTOR AMPLIFIER

NAME THE DEVICE THAT PROVIDES AN INCREASE IN CURRENT, VOLTAGE, OR POWER OF A SIGNAL WITHOUT APPRECIABLY ALTERING THE ORIGINAL SIGNAL.

BESIDES ELIMINATING THE EMITTER-BASE BATTERY, SPECIFY THE OTHER ADVANTAGES THAT DIFFERENT BIASING METHODS OFFER.

STATE THE PRIMARY DIFFERENCE BETWEEN THE NPN AND PNP AMPLIFIERS.

IDENTIFY THE MOST UNSTABLE BIASING METHOD.

SPECIFY THE MOST WIDELY USED COMBINATION-BIAS SYSTEM.

SPECIFY THE AMPLIFIER CLASS OF OPERATION THAT ALLOWS COLLECTOR CURRENT TO FLOW DURING THE COMPLETE CYCLE OF THE INPUT.

LIST TWO PRIMARY ITEMS THAT DETERMINE THE CLASS OPERATION OF AN AMPLIFIER.

In the preceding pages we explained the internal workings of the transistor and introduced new terms, such as emitter, base, and collector. Since you should be familiar by now with all of these new terms and with the internal operation of the transistor, we will move on to the basic transistor amplifier.

To understand the overall operation of the transistor amplifier, you must only consider the current in and out of the transistor and through the various components in the circuit. Therefore, from this point on, only the schematic symbol for the transistor will be used in the illustrations, and rather than thinking about majority and minority carriers, we will now start thinking in terms of emitter, base, and collector current.

Before going into the basic transistor amplifier, there are two terms you should be familiar with: AMPLIFICATION and AMPLIFIER. Amplification is the process of increasing the strength of a SIGNAL. A signal is a general term used to refer to any particular current, voltage, or power in a circuit. An amplifier is the device that provides amplification (the increase in current, voltage, or power of a signal) without appreciably altering the original signal.

Transistors are frequently used as amplifiers. Some transistor circuits are CURRENT amplifiers, with a small load resistance; other circuits are designed for VOLTAGE amplification and have a high load resistance; others amplify POWER.

Now take a look at the NPN version of the basic transistor amplifier in figure 2-15 and let's see just how it works.

So far in this discussion, a separate battery has been used to provide the necessary forward-bias voltage. Although a separate battery has been used in the past for convenience, it is not practical to use a battery for emitter-base bias. For instance, it would take a battery slightly over .2 volts to properly forward bias a germanium transistor while a similar silicon transistor would require a voltage slightly over .6 volts. However, common batteries do not have such voltage values. Also, since bias voltages are quite critical and must be held within a few tenths of one volt, it is easier to work with bias currents flowing through resistors of high ohmic values than with batteries.

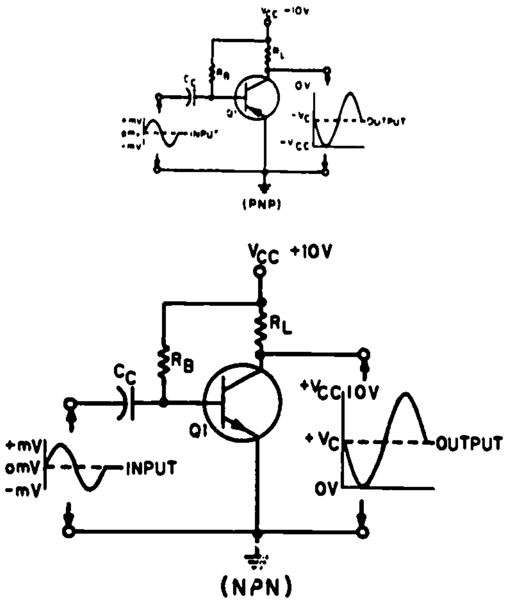


Fig 2-15. The basic transistor amplifier.

By inserting one or more resistors in a circuit, different methods of biasing may be achieved and the emitter-base battery eliminated. In addition to eliminating the battery, some of these biasing methods compensate for slight variations in transistor characteristics and changes in transistor conduction resulting from temperature irregularities. Notice in figure 2-15 that the emitter-base battery has been eliminated and the bias resistor R_B has been inserted between the collector and the base. Resistor R_B provides the necessary forward bias for the emitter-base junction. Current flows in the emitter-base bias circuit from ground to the emitter, out the base lead, and through R_B to V_{CC}. Since the current in the base circuit is very small (a few hundred microamperes) and the forward resistance of the transistor is low, only a few tenths of a volt of positive bias will be felt on the base of the transistor. However, this is enough voltage on the base, along with ground on the emitter and the large positive voltage on the collector, to properly bias the transistor.

With Q1 properly biased, direct current flows continuously, with or without an input signal throughout the entire circuit. The direct current flowing through the circuit develops more than just base bias; it also develops the collector voltage (V_C) as it flows through Q1 and R_L. Notice the collector voltage on the output graph. Since it is presented into the circuit without an input signal, then the output signal starts at the V_C level and either increases or decreases. These d.c. voltages and currents that exist in the circuit prior to the application of signal are known as <u>QUIESCENT voltages and currents</u> (the quiescent state of the circuit).

Resistor R_L, the collector load resistor, is placed in the circuit to keep the full effect of the collector supply voltage off the collector. This permits the collector voltage (V_C) to change with an input signal, which in turn allows the transistor to amplify voltage. Without R_L in the circuit, the voltage on the collector would always be equal to V_{CC}.

The coupling capacitor (C_C) is another new addition to the transistor circuit. It is used to pass the a.c. input signal and block the d.c. voltage from the preceding circuit. This prevents d.c. in the circuitry on the left of the coupling capacitor from affecting the bias on Q1. The coupling capacitor also blocks the bias of Q1 from reaching the input signal source.

The input to the amplifier is a sine wave that varies a few millivolts above and below zero. It is introduced into the circuit by the coupling capacitor and is applied between the base and emitter. As the input signal goes positive, the voltage across the emitter-base junction becomes more positive. This, in effect, increases forward bias which causes base current to increase at the same rate as that of the input sine wave. Emitter and collector currents also increase but much more than the base current. With an increase in collector current, more voltage is developed across R_L. Since the voltage across R_L and the voltage across Q1 (collector to emitter) must add up to V_{CC}, an increase in voltage across R_L results in an equal decrease in voltage across Q1. Therefore, the output voltage from the amplifier, taken at the collector of Q1 with respect to the emitter, is a negative alternation of voltage that is larger than the input but has the same sine wave characteristics.

During the negative alternation of the input, the input signal opposes the forward bias. This action decreases base current which results in a decrease in both emitter and collector currents. The decrease in current through R_L decreases its voltage drop and causes the voltage across the transistor to rise along with the output voltage. Therefore, the output for the negative alternation of the input is a positive alternation of voltage that is larger than the input but has the same sine wave characteristics. By examining both input and output signals for one complete alternation of the input, we can see that the output of the amplifier is an exact reproduction of the input except for the reversal in polarity and the increased amplitude (a few millivolts as compared to a few volts).

The PNP version of this amplifier is shown in the upper right-hand corner of the figure 2-15. The primary difference between the NPN and PNP amplifier is the polarity of the source voltage. With a negative V_{CC}, the PNP base voltage is slightly negative with respect to ground, which provides the necessary forward bias condition between the emitter and base.

When the PNP input signal goes positive, it opposes the forward bias of the transistor. This action cancels some of the negative voltage across the emitter-base junction which reduces the current through the transistor. Therefore, the voltage across the load resistor decreases, and the voltage across the transistor increases. Since V_{CC} is negative, the voltage on the collector (V_C) goes in a negative direction (as shown on the output graph) toward - V_{CC} (For example, from -5 volts to - 7 volts). Thus, the output is a negative alternation of voltage that varies at the same rate as the sine wave input but is opposite in polarity and has a much larger amplitude. During the negative alternation of the input signal, the transistor current increases because the input voltage aids the forward bias. Therefore, the voltage across R_L increases, and consequently, the voltage across the transistor decreases or goes in a positive direction (For example: from - 5 volts to - 3 volts). This action results in a positive output voltage which has the same characteristics as the input except that it has been amplified and the polarity is reversed.

In summary, the input signals in the preceding circuits were amplified because the small change in base current caused a large change in collector current, and by placing resistor R_L in series with the collector, voltage amplification was achieved.

One of the basic problems with transistor amplifiers is establishing and maintaining the proper values of quiescent current and voltage in the circuit. This is accomplished by selecting the proper circuit-biasing conditions and ensuring these conditions are maintained despite the variations in ambient (surrounding) temperature, which cause changes in amplification and even distortion (an unwanted change in a signal). Thus, a need arises for a method to properly bias the transistor amplifier and at the same time stabilize its d.c. operation point (the no signal values of collector voltage and collector current). As mentioned earlier, various biasing methods can be used to accomplish both of these functions. Although there are numerous biasing methods, only three basic types will be considered.

The first biasing method, called BASE CURRENT BIAS or sometimes FIXED BIAS, was used in figure 2-15. As you recall, it consisted basically of a resistor (R_B) connected between the collector supply voltage and the base. Unfortunately, this simple arrangement is quite thermally unstable. If the temperature of the transistor rises for any reason (due to a rise in ambient temperature or due to current flow through it), collector current will increase. This increase in current also causes the d.c. operating point, sometimes called the quiescent or static point, to move from its desired position (level). This reaction to temperature is undesirable because it affects amplifier gain (the number of times of amplification) and could result in distortion, as you will see later in this discussion.

A better method of biasing is obtained by inserting the bias resistor directly between the base and collector, as shown in figure 2-16. By tying the collector to the base in this manner, feed-back voltage can be fed from the collector to the base to develop forward bias. This arrangement is called SELF-BIAS. Now, if an increase of temperature causes an increase in collector current, the collector voltage (V_C) will fall due to the increase of voltage produced across the load resistor (R_L). This drop in V_C will be fed back to the base and will result in a decrease in the base current. The decrease in base current will oppose the original increase in collector current and tend to stabilize it. The exact opposite effect is produced when the collector current decreases.

Self-bias has two small drawbacks: (1) It is only partially effective and, therefore, is only used where moderate changes in ambient temperature are expected; (2) it reduces amplification since the signal on the collector also affects the base voltage. This is because the collector and base signals for this particular amplifier configuration are 180 degrees out of phase (opposite in polarity) and the part of the collector signal that is fed back to the base cancels some of the input signal. This process of returning a part of the output back to its input is known as <u>DEGENERATION</u> or <u>NEGATIVE FEEDBACK</u>. Sometimes degeneration is desired to prevent amplitude distortion (an output signal that fails to follow the input exactly) and self-bias may be used for this purpose.

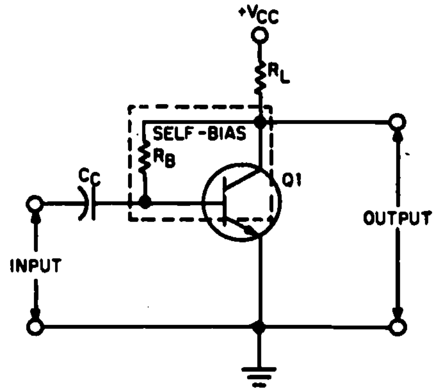


Fig 2-16. A basic transistor amplifier with self-bias.

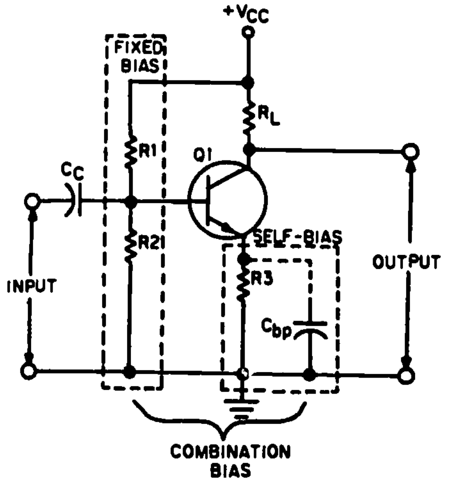


Fig 2-17. A basic transistor amplifier with combination bias.

A combination of fixed and self-bias can be used to improve stability and at the same time overcome some of the disadvantages of the other two biasing methods. One of the most widely used combination-bias systems is the voltage-divider type shown in figure 2-17. Fixed bias is provided in this circuit by the voltage-divider network consisting of R1, R2, and the collector supply voltage (V_{CC}). The d.c. current flowing through the voltage-divider network biases the base positive with respect to the emitter. Resistor R3, which is connected in series with the emitter, provides the emitter with self-bias. Should I_E increase, the voltage drop across R_3 would also increase, reducing V_C. This reaction to an increase in I_E by R_3 is another form of degeneration, which results in less output from the amplifier. However, to provide long-term or d.c. thermal stability and at the same time allow minimal a.c. signal degeneration, the bypass capacitor (C_{bp}) is placed across R3. If C_{bp} is large enough, rapid signal variations will not change its charge materially and no degeneration of the signal will occur.

In summary, the fixed-bias resistors, R_1 and R_2, tend to keep the base bias constant while the emitter bias changes with emitter conduction. This action greatly improves thermal stability and at the same time maintains the correct operating point of the transistor.

In the previous discussions, we assumed that for every portion of the input signal there was an output from the amplifier. This is not always the case with amplifiers. It may be desirable to have the transistor conducting for only a portion of the input signal. The portion of the input for which there is an output determines the class of operation of the amplifier. There are four classes of amplifier operations. They are class A, class AB, class B, and class C.

Class A amplifiers are biased so that variations in input signal polarities occur within the limits of CUTOFF and SATURATION. In a PNP transistor, for example, if the base becomes positive with respect to the emitter, holes will be repelled at the PN junction and no current can flow in the collector circuit. This condition is known as cutoff. Saturation occurs when the base becomes so negative with respect to the emitter that changes in the signal are not reflected in collector-current flow.

Biasing an amplifier in this manner places the d.c. operating point between cutoff and saturation and allows collector current to flow during the complete cycle (360 degrees) of the input signal, thus providing an output which is a replica of the input. Figure 2-15 is an example of a class A amplifier. Although the output from this amplifier is 180 degrees out of phase with the input, the output current still flows for the complete duration of the input.

The class A operated amplifier is used as an audio-and radio-frequency amplifier in radio, radar, and sound systems, just to mention a few examples.

For a comparison of output signals for the different amplifier classes of operation, refer to figure 2-18 during the following discussion.

Amplifiers designed for class AB operation are biased so that collector current is zero (cutoff) for a portion of one alternation of the input signal. This is accomplished by making the forward-bias voltage less than part value of the input signal. By doing this, the base-emitter junction will be reverse biased during one alternation for the amount of time that the input signal voltage. Therefore, collector current will flow for more than 180 degrees but less than 360 degrees of the input signal, as shown in figure 2-18(B). As compared to the class A amplifier, the d.c. operating point for the class AB amplifier is closer to cutoff.

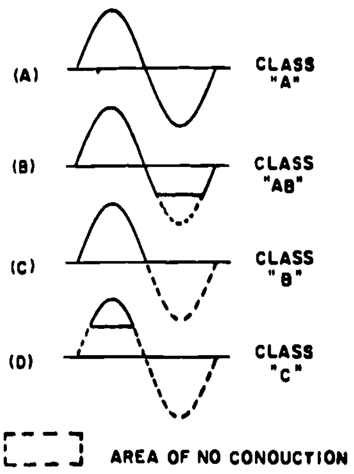


Fig 2-18. A comparison of output signals for the different amplifier classes of operation.

The class AB operated amplifier is commonly used as a push-pull amplifier to overcome a side effect of class B operation called crossover distortion.

Amplifiers biased so that collector current is cut off during one-half of the input signal are classified class B. The d.c. operating point for this class of amplifier is set up so that base current is zero with no input signal. When a signal is applied, one half cycle will forward bias the base-emitter junction and I_C will flow. The other half cycle will reverse bias the base-emitter junction and I_C will be cut off. Thus, for class B operation, collector current will flow for approximately 180 degrees (half) of the input signal, as shown in figure 2-18(C).

The class B operated amplifier is used extensively for audio amplifiers that require high-power outputs. It is also used as the driver-and power-amplifier stages of transmitters.

In class C operation, collector current flows for less than one half cycle of the input signal, as show in figure 2-18(D). The class C operation is achieved by reverse biasing the emitter-base junction which sets the d.c. operating point below cutoff and allows only the portion of the input signal that overcomes the reverse bias to cause collector current flow.

The class C operated amplifier is used as a radio-frequency amplifier in transmitters.

From the previous discussion, you can conclude that two primary items determine the class of operation of an amplifier--(1) the amount of bias and (2) the amplitude of the input signal. With a given input signal and bias level, you can change the operation of an amplifier from class A to class B just by removing forward bias. Also, a class A amplifier can be changed to class AB by increasing the input signal amplitude. However, if an input signal amplitude is increased to the point that the transistor goes into saturation and cutoff, it is then called an OVER-DRIVEN amplifier.

There are two terms used in conjunction with amplifiers that you should be familiar with--FIDELITY and EFFICIENCY. Fidelity is the faithful reproduction of a signal. In other words, if the output of an amplifier is just like the input except in amplitude, the amplifier has a high degree of fidelity. The opposite of fidelity is a term we mentioned earlier--distortion. Therefore, a circuit that has high fidelity has low distortion. In conclusion, a class A amplifier has a high degree of fidelity. A class AB amplifier has less fidelity, and class B and C amplifiers have low or "poor" fidelity.

The efficiency of an amplifier refers to the ratio of output-signal power compared to the total input power. An amplifier has two input power sources: one from the signal, and one from the power supply. Since every device takes power to operate, an amplifier that operates for 360 degrees of the input signal uses more power than if operated for 180 degrees of the input signal. By using more power, an amplifier has less power available for the output signal; thus the efficiency of the amplifier is low. This is the case with the class A amplifier. It operates for 360 degrees of the input signal and requires a relatively large input from the power supply. Even with no input signal, the class A amplifier still uses power from the power supply. Therefore, the output from the class A amplifier is relatively small compared to the total input power. This results in low efficiency which is acceptable in class A amplifiers because they are used where efficiency in not as important as fidelity.

Class AB amplifiers are biased so that collector current is cut off for a portion of one alternation of the input which results in less total input power than the class A amplifier. This leads to better efficiency.

Class B amplifiers are biased with little or no collector current at the d.c. operating point. With no input signal, there is little wasted power. Therefore, the efficiency of class B amplifiers is higher still.

The efficiency of class C is the highest of the four classes of amplifier operations.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is the name of the device that provides an increase in current, voltage, or power of a signal without appreciably altering the orginal signal?

2. Besides eliminating the emitter base battery, what other advantages can different biasing methods offer?

3. In the basic transistor amplifier discussed earlier, what is the relationship between the polarity of the input and output signals?

4. What is the primary difference between the NPN and PNP amplifiers?

5. Which biasing method is the most unstable?

6. What type of bias is used where only moderate changes in ambient temperature are expected?

7. When is degeneration tolerable in an amplifier?

8. What is the most widely used combination-bias system?

9. What amplifier class of operation allows collector current to flow during the complete cycle of the input?

10. What is the name of the term used to describe the condition in a transistor when the emitter-base junction has zero bias or is reverse biased and there is no collector current?

11. List the two primary items that determine the class of operation of an amplifier.

a. ____________________

b. ____________________

12. What amplifier class of operation is the most inefficient but has the least distortion?

____

Work Unit 2-4. TRANSISTOR CONFIGURATIONS

SPECIFY THE TRANSISTOR CONFIGURATION THAT PROVIDES A PHASE REVERSAL BETWEEN THE INPUT AND OUTPUT SIGNALS.

SPECIFY THE TRANSISTOR CONFIGURATION THAT HAS A CURRENT GAIN OF LESS THAN 1.

STATE THE INPUT CURRENT IN THE COMMON-EMITTER CIRCUIT.

STATE THE FORMULA FOR GAMMA.

A transistor may be connected in any one of three basic configurations (figure 2-19): common emitter (CE), common base (CB), and common collector (CC). The term "common" is used to denote the element that is common to both input and output circuits. Because the common element is often grounded, these configurations are frequently referred to as grounded emitter, grounded base, and grounded collector.

Each configuration, as you will see later, has particular characteristics that make it suitable for specific applications. An easy way to identify a specific transistor configuration is to follow three simple steps:

(a) Identify the element (emitter, base of collector) to which the input signal is applied.

(b) Identify the element (emitter base, or collector) from which the output signal is taken.

(c) The remaining element is the common element, and gives the configuration its name.

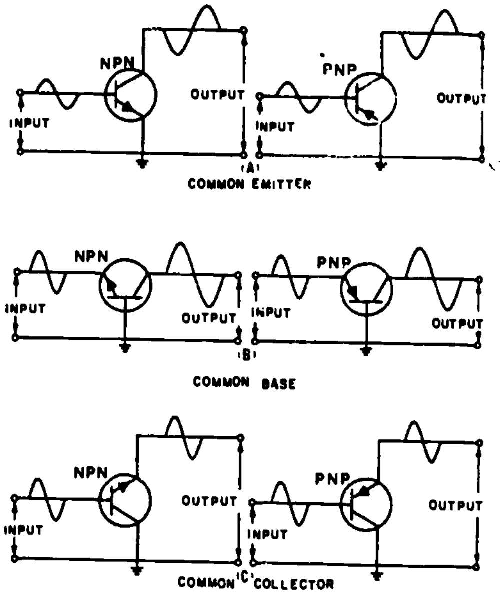


Fig 2-19. Transistor configurations.

Therefore, by applying these three simple steps to the circuit in figure 2-15, we can conclude that this circuit is more than just a basic transistor amplifier. It is a common-emitter amplifier.

The common-emitter configuration (CE) shown in figure 2-19(A) is the arrangement most frequently used in practical amplifier circuits, since it provides good voltage, current, and power gain. The common emitter also has a somewhat low input resistance (500 ohms--1500 ohms), because the input is applied to the forward-biased junction, and a moderately high output resistance (30) kilohms--50 kilohms or more), because the output is taken off the reverse-biased junction. Since the input signal is applied to the base-emitter circuit and the output is taken from the collector-emitter circuit, the emitter is the element common to both input and output.

Since you have already covered what you now know to be a common-emitter amplifier (figure 2-15), let's take a few minutes and review its operation, using the PNP common-emitter configuration shown in figure 2-19(A).

When a transistor is connected in a common-emitter configuration, the input signal is injected between the base and emitter, which is a low resistance, low-current circuit. As the input signal swings positive, it also causes the base to swing positive with respect to the emitter. This action decreases forward bias which reduces collector current (I_C) and increases collector voltage (making V_C more negative). During the negative alternation of the input signal, the base is driven more negative with respect to the emitter. This increases forward bias and allows more current carriers to be released from the emitter which results in an increase in collector current and a decrease in collector voltage (making V_C less negative or swing in a positive direction). The collector current which flows through the high resistance reverse-biased junction also flows through a high resistance load (not shown), resulting in a high level of amplification.

Since the input signal to the common emitter goes positive when the output goes negative, the two signals (input and output) are 180 degrees out of phase. The common-emitter circuit is the only configuration which provides a phase reversal.

The common emitter is the most popular of the three transistor configurations because it has the best combination of current and voltage gain. The term GAIN is used to describe the amplification capabilities of the amplifier. It is basically a ratio of output divided by input. Each transistor configuration gives a different value of gain even though the same transistor is used. The transistor configuration used is a matter of design consideration. However, as a technician you will become interested in this output versus input ratio (gain) in order to determine whether or not the transistor is working properly in the circuit.

The current gain in the common-emitter circuit is called BETA (B). Beta is the relationship of collector current (output current) to base current (input current). To calculate beta, use the following formula:

$B = \frac{\Delta I_C}{\Delta I_B}$ (Δ is the Greek letter delta, it is used to indicate a small change)

For example, if the input current (I_B) in a common emitter changes from 75   A to 100   A and the output current (I_C) changes from 1.5 mA to 2.6 mA, the current again (B) would be 44.

$$B = \frac{I_C}{I_B} = \frac{1.1 \times 10^{-3}}{25 \times 10^{-6}} = 44$$

This simply means that a change in base current produces a change in collector current which is 44 times as large.

You may also see the term h_{fe} used in place of B. The terms h_{fe} and B are equivalent and may be used interchangeably. This is because "h_{fe}" means:

h = hybrid (meaning mixture)
f = forward current transfer ratio
e = common emitter configuration

The resistance gain of the common emitter can be found in a method similar to the one used for finding beta:

$$R = \frac{R\ out}{R\ in}$$

Once the resistance gain is known, the voltage gain is easy to calculate since it is equal to the current gain (B) multiplied by the resistance gain ($E = BR$). Therefore, the power gain is equal to the voltage gain multiplied by the current gain B ($P = BE$).

The common-base configuration (CB) shown in figure 2-19(B) is mainly used for impedance matching, since it has a low input resistance (30 ohms—160 ohms) and a high output resistance (250 kilohms—550 kilohms). However, two factors limit its usefulness in some circuit applications: (1) its low input resistance and (2) its current gain of less than 1. Since the CB configuration will give voltage amplification, there are some additional applications, which require both a low input resistance and voltage amplification, that could use a circuit configuration of this type; for example, some microphone amplifiers.

In the common-base configuration, the input signal is applied to the emitter, the output is taken from the collector, and the base is the element common to both input and output. Since the input is applied to the emitter, it causes the emitter-base junction to react in the same manner as it did in the common-emitter circuit. For example, an input that aids the bias will increase transistor current, and one that opposes the bias will decrease transistor current.

Unlike the common-emitter circuit, the input and output signals in the common-base circuit are in phase. To illustrate this point, assume the input to the PNP version of the common-base circuit in figure 2-19(B) is positive. The signal adds to the forward bias, since it is applied to the emitter, causing the collector current to increase. This increase in I_C results in a greater voltage drop across the load resistor R_L (not shown), thus lowering the collector voltage V_C. The collector voltage, in becoming less negative, is swinging in a positive direction and is, therefore, in phase with the incoming positive signal.

The current gain in the common-base circuit is calculated in a method similar to that of the common emitter except that the input current is I_E not I_B and the term ALPHA (a) is used in place of beta for gain. Alpha is the relationship of collector current (output current) to emitter current (input current). Alpha is calculated using the formula:

$$a = \frac{\Delta I_C}{\Delta I_E}$$

For example, if the input current (I_e) in a common base changes from 1 mA to 3m A and the output current (I_C) changes from 1 mA to 2.8 mA, then the current gain (a) would be 0.90 or:

$$a = \frac{\Delta I_C}{\Delta I_E}$$

$$= \frac{1.8 \times 10^{-3}}{2 \times 10^{-3}}$$

$$= 0.90$$

This is a current gain of less than 1.

Since part of the emitter current flows into the base and does not appear as collector current, then collector current will always be less than the emitter current that causes it. (Remember, $I_E = I_B + I_C$) Therefore, ALPHA is ALWAYS LESS THAN ONE FOR A COMMON-BASE CONFIGURATION.

Another term for "a" is h_{fb}. These terms ("a" and h_{fb}) are equivalent and may be used interchangeably. The meaning for the term h_{fb} is derived in the same manner as the term h_{fe} mentioned earlier, except that the last letter "e" has been replaced with "b" to stand for common-base configuration.

Many transistor manuals and data sheets only list transistor current gain characteristics in terms of B or h_{fe}. To find alpha (a) when given beta (B), use the following formula to convert B to a for use with the common-base configuration:

$$a = \frac{B}{B + 1}$$

To calculate the other gains (voltage and power) in the common-base configuration when the current gain (a) is known, follow the procedures described earlier under the common-emitter section.

The common-collector configuration (CC) shown in figure 2-19(C) is used mostly for impedance matching. It is also used as a current driver, due to its substantial current gain. It is particularly useful in switching circuitry, since it has the ability to pass signals in either direction (bilateral operation).

In the common-collector circuit, the input signal is applied to the base, the output is taken from the emitter, and the collector is the element common to both input and output. The common collector is equivalent to our old friend the electron-tube cathode follower. Both have high input and low output resistance. The input resistance for the common collector ranges from 2 kilohms to 500 kilohms and the output resistance varies from 50 ohms to 1500 ohms. The current gain is higher than that in the common emitter, but it has a lower power gain than either the common base or common emitter. Like the common base, the output signal from the common collector is in phase with the input signal. The common collector is also referred to as an emitter-follower because the output developed on the emitter follows the input signal applied to the base.

Transistor action in the common collector is similar to the operation explained for the common base, except that the current gain is not based on the emitter-to-collector current ratio, alpha (a). Instead, it is based on the emitter-to-base current ratio called GAMMA (v), because the output is taken of the emitter. Since a small change in base current controls a large change in emitter current, it is still possible to obtain high current gain in the common collector. However, since the emitter current gain is offset by the low output resistance, the voltage gain is always less than 1 (unity), exactly as in the electron-tube cathode follower.

The common-collector current gain, gamma (v), is defined as

$$V = \frac{I_E}{I_B}$$

and is related to collector-to-base current gain, beta (B), of the common-emitter circuit by the formula:

$$v = B + 1$$

Since a given transistor may be connected in any of three basic configurations, then there is a define relationship, as pointed out earlier, between alpha, beta, and gamma. These relationships are listed again for your convenience:

$$\text{Alpha (a)} = \frac{B}{B + 1}$$

$$\text{Beta (B)} = \frac{a}{1 - a}$$

$$\text{Gamma (v)} = B + 1$$

Take, for example, a transistor that is listed on a manufacturer's data sheet as having an alpha of 0.90, but we wish to use it in a common-emitter configuration. This means we must find beta. The calculations are:

$$B = \frac{a}{1 - a} = \frac{0.90}{1 - 0.90} = \frac{0.90}{0.1} = 9$$

Therefore, a change in base current in this transistor will produce a change in collector current that will be 9 times as large.

If we wish to use this same transistor in a common collector, we can find gamma by:

$$\text{Gamma} = B + 1 = 9 + 1 = 10$$

To summarize the properties of the three transistor configurations, a comparison chart is provided in table 2-1 for your convenience.

Table 2-1.

Amplifier Type	Common Base	Common Emitter	Common Collector
Input/output Phase Relationship	0°	180°	0°
Voltage Gain	HIGH	MEDIUM	LOW
Current Gain	LOW (a)	MEDIUM (B)	HIGH (2)
Power Gain	LOW	HIGH	MEDIUM
Input Resistance	LOW	MEDIUM	HIGH
Output Resistance	HIGH	MEDIUM	LOW

Now that we have analyzed the basic transistor amplifier in terms of bias, class of operation, and circuit configuration, let's apply what has been covered to figure 2-15. A reproduction of figure 2-15 is shown below for your convenience.

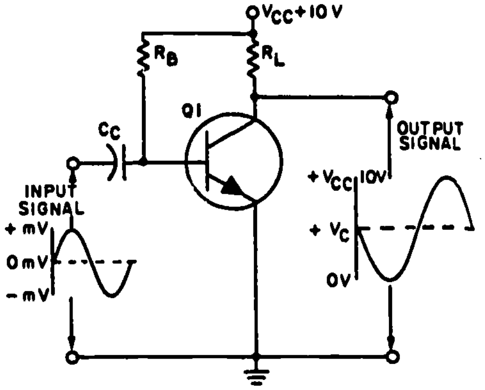


Fig 2-15. The basic transistor amplifier.

This illustration is not just the basic transistor amplifier shown earlier in figure 2-15 but a class A amplifier configured as a common emitter using fixed bias. From this, you should be able to conclude the following:

Because of its fixed bias, the amplifier is thermally unstable.

Because of its class A operation, the amplifier has low efficiency but good fidelity.

Because it is configured as a common emitter, the amplifier has good voltage, current, and power gain.

In conclusion, the type of bias, class of operation, and circuit configuration are all clues to the function and possible application of the amplifier.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What are the three transistor configurations?

a. ______________________________

b. ______________________________

c. ______________________________

2. Which transistor configuration provides a phase reversal between the input and output signals?

____

3. What is the input current in the common-emitter circuit?

____

4. What is the current gain in a common-base circuit called?

____

5. Which transistor configuration has a current gain of less than 1?

____

6. What is the formula for GAMMA?

____

Work Unit 2-5. TRANSISTOR SPECIFICATION, AND IDENTIFICATION

LIST THREE ITEMS OF INFORMATION NORMALLY INCLUDED IN THE GENERAL DESCRIPTION SECTION OF A SPECIFICATION SHEET FOR A TRANSISTOR.

SPECIFY WHAT THE NUMBER "2" (BEFORE THE LETTER "N") INDICATES IN THE JAN MARKING SCHEME.

Transistors, like electron tubes, are available in a large variety of shapes and sizes, each with its own unique characteristics. The characteristics for each of these transistors are usually presented on SPECIFICATION SHEETS or they may be included in transistor manuals. Although many properties of a transistor could be specified on these sheets, manufacturers list only some of them. The specifications listed vary with different manufacturers, the type of transistor, and the application of the transistor. The specifications usually cover the following items.

a A general description of the transistor which includes the following information:

(1) The kind of transistor. This covers the material used, such as germanium or silicon; the type of transistor (NPN or PNP); and the construction, grown, or diffused junction, etc.).

(2) Some of the common applications for the transistor, such as audio amplifier, oscillator, of amplifier, etc.

(3) General sales features, such as size and packaging (mechanical data).

b. The "Absolute Maximum Ratings" of the transistor are the direct voltage and current values that if exceeded in operation may result in transistor failure. Maximum ratings usually include collector-to-base voltage, emitter-to-base voltage, collector current, emitter current, and collector power dissipation.

c. The typical operating values of the transistor. These values are presented only as a guide. The values vary widely, are dependent upon operating voltage, and also upon which element is common in the circuit. The values listed may include collector-emitter voltage, collector current, input resistance, load resistance, current-transfer ratio (another name for alpha or beta), and collector cutoff current, which is leakage current from collector to base when no emitter current is applied. Transistor characteristic curves may also be included in this section. A transistor characteristic curve is a graph plotting the relationship between currents and voltage in a circuit. More than one curve on a graph is called a "family of curves."

Transistors can be identified by a Joint Army-Navy (JAN) designation printed directly on the case of the transistor. The marking scheme explained earlier for diodes is also used for transistor identification. The first number indicates the number of junctions. The letter "N" following the first number tells us that the component is a semiconductor. And, the 2- or 3-digit number following the N is the manufacturer's identification number. If the last number is followed by a letter, it indicates an improved version of the device. For example, a semiconductor designated as type 2N130A signifies a three-element transistor of semiconductor material that is an improved version of type 130:

2	N	130	A
NUMBER OF JUNCTIONS (TRANSISTOR)	SEMI-CONDUCTOR	IDENTIFICATION NUMBER	FIRST MODIFICATION

You may also find other markings on transistors which do not relate to the JAN marking system. These markings are manufacturer's identifications and may not conform to a standardized system. If in doubt, always replace a transistor with one having identical markings. To ensure that an identical replacement or a correct substitute is used, consult an equipment or transistor manual for specifications on the transistor.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. List three items of information normally included in the general description section of a specification sheet for a transistor.

   a. ______________________________

   b. ______________________________

   c. ______________________________

2. What does the number "2" (before the letter "N") indicate in the JAN marking scheme?

   ____

   ____

Work Unit 2-6. TRANSISTOR MAINTENANCE.

SPECIFY THE GREATEST DANGER TO A TRANSISTOR.

IDENTIFY THE METHODS FOR CHECKING TRANSISTORS THAT ARE CUMBERSOME WHEN MORE THAN ONE TRANSISTOR IS BAD IN A CIRCUIT.

STATE THE SAFETY PRECAUTION THAT MUST BE TAKEN BEFORE REPLACING A TRANSISTOR.

STATE HOW THE COLLECTOR LEAD IS IDENTIFIED ON AN OVAL-SHAPED TRANSISTOR.

STATE THE TWO MOST IMPORTANT PARAMETERS USED FOR TESTING A TRANSISTOR.

Transistors, unlike electron tubes, are very rugged and are expected to be relatively trouble free. Encapsulation and conformal coating techniques now in use promise extremely long life expectancies. In theory, a transistor should last indefinitely. However, if transistors are subjected to current overloads, the junctions will be damaged or even destroyed. In addition, the application of excessively high operating voltages can damage or destroy the junctions through arc-over or excessive reverse currents. One of the greatest dangers to the transistor is heat, which will cause excessive current flow and eventual destruction of the transistor.

In order to determine if a transistor is good or bad, you can check it with an ohmmeter or a transistor tester. In many cases, you can substitute a transistor known to be good for one that is questionable and thus determine the condition of a suspected transistor. This method of testing is highly accurate and sometimes the quickest, but it should be used only after you make certain that there are no circuit defects that might damage the replacement transistor. If more than one defective transistor is present in the equipment where the trouble has been localized, this testing method becomes cumbersome, as several transistors may have to be replaced before the trouble is corrected. To determine which stages failed and which transistors are not defective, all the removed transistors must be tested. This test can be made by using a standard ohmmeter, transistor tester, or by observing whether the equipment operates correctly as each of the removed transistors is reinserted into the equipment. A word of caution--indiscriminate substitution of transistors in critical circuits should be avoided.

When transistors are soldered into equipment, substitution is not practicable; it is generally desirable to test these transistors in their circuits. Transistors, although generally more rugged mechanically than electron tubes, are susceptible to damage by electrical overloads, heat, humidity, and radiation. Damage of this nature often occurs during transistor servicing by applying the incorrect polarity voltage to the collector circuit or excessive voltage to the input circuit. Careless soldering techniques that overheat the transistor have also been known to cause considerable damage. One of the most frequent causes of damage to a transistor is the electrostatic discharge from the human body when the device is handled. You may avoid such damage before starting repairs by discharging the static electricity from your body to the chassis containing the transistor. You can do this by simply touching the chassis. Thus, the electricity will be transferred from your body to the chassis before you handle the transistor.

To prevent transistor damage and avoid electrical shock, you should observe the following precautions when you are working with transistorized equipment:

a. Test equipment and soldering irons should be checked to make certain that there is no leakage current from the power source. If leakage current is detected, isolation transformers should be used.

b. Always connect a ground between test equipment and circuit before attempting to inject or monitor a signal.

c. Ensure test voltage do not exceed maximum allowable voltage for circuit components and transistors. Also, never connect test equipment outputs directly to a transistor circuit.

d. Ohmmeter ranges which require a current of more than one milliampere in the test circuit should not be used for testing transistor.

e. Battery eliminators should not be used to furnish power for transistor equipment because they have poor voltage regulation and, possibly, high-ripple voltage.

f. The heat applied to a transistor, when soldered connections are required, should be kept to a minimum by using a low-wattage soldering iron and heat shunts, such as long-nose pliers, on the transistor leads.

g. When it becomes necessary to replace transistors, never pry transistors to loosen them from printed circuit boards.

h. All circuits should be checked for defects before replacing a transistor.

i. The power must be removed from the equipment before replacing a transistor.

j. Using conventional test probes on equipment with closely spaced parts often causes accidental shorts between adjacent terminals. These shorts rarely cause damage to an electron tube but may ruin a transistor. To prevent these shorts, the probes can be covered with insulation, except for a very short length of the tips.

Transistor lead identification plays an important part in transistor maintenance, because before a transistor can be test or replaced, its leads or terminals must be identified. Since there is no standard method of identifying transistor leads, it is quite possible to mistake one lead for another. Therefore, when you are replacing a transistor, you should pay close attention to how the transistor is mounted, particularly to those transistors that are soldered in, so that you do not make a mistake when you are installing the new transistor. When you are testing or replacing a transistor, if you have any doubts about which lead is which, consult the equipment manual or a transistor manual that shows the specifications for the transistor being used.

There are, however, some typical lead identification schemes that will be very helpful in transistor troubleshooting. These schemes are shown in figure 2-20. In the case of the oval shaped transistors shown in view (A), the collector lead is identified by a wide space between it and the base lead. The lead farthest from the collector, in line, is the emitter lead. When the leads are evenly spaced and in line, as shown in view (B), a colored dot, usually red, indicates the collector. If the transistor is round, as in view (C), a red line indicates the collector, and the emitter lead is the shortest lead. In view (D) the leads are in a triangular arrangement which is offset from the center of the transistor. The lead opposite the blank quadrant in this scheme is the base lead. When viewed from the bottom, the collector is the first lead clockwise from the base. The leads in view (E) are arranged in the same manner as those in view (D) except that a tap is used to identify the leads. When viewed from the bottom in a clockwise direction, first lead following the tab is the emitter, followed by the base and collector.

In a conventional power transistor as shown in views (F) and (G), the collector lead is usually connected to the mounting base. For further identification, the base lead in view (F) and (G) are identified by viewing the transistor from the bottom in a clockwise direction (with mounting holes occupying 3 o'clock and 9 o'clock positions), the emitter lead will be either at the 5 o'clock or 11 o'clock position. The other lead is the base lead.

There are several different ways of testing transistors. They can be tested while in the circuit, by the substitution method mentioned, or with a transistor tester or ohmmeter.

Transistor testers are nothing more than the solid-state equivalent of electron-tube testers (although they do not operate on the same principle). With most transistor testers it is possible to test the transistor in or out of the circuit.

Since it is impractical to cover all the different types of transistor testers and since each tester comes with its own operator's manual, we will move on to something you will use more frequently for testing transistors--the ohmmeter.

There are four basic tests required for transistor in practical troubleshooting: gain, leakage, breakdown, and switching time. For maintenance and repair, however, it is usually not necessary to check all of these parameters. A check of two or three parameters is usually sufficient to determine whether a transistor needs to be replaced. Two of the most important parameters used for testing are gain and leakage. The following paragraphs will give you a good idea of how to use the ohmmeter to check for transistor gain and leakage.

TRANSISTOR GAIN TEST -- A basic transistor gain test can be made using an ohmmeter and a simple test circuit. The test circuit can be made with just a couple of resistors and a switch, as shown in figure 2-21. The principle behind the test lies in the fact that little or no current will flow in a transistor between emitter and collector until the emitter-base junction is forward biased. The only precaution you should observe is with the ohmmeter. Any internal battery may be used in the meter provided it does not exceed the maximum collector-emitter breakdown voltage.

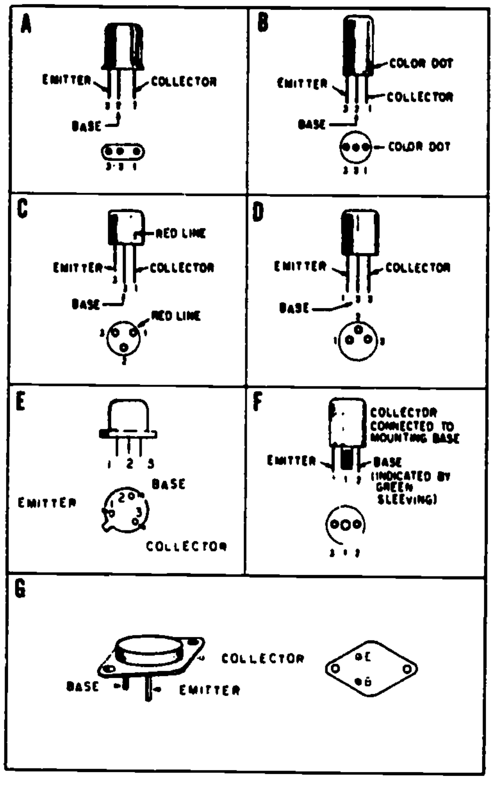


Fig 2-20. Transistor lead identification.

With the switch in figure 2-21 in the open position as shown, no voltage is applied to the PNP transistor's base, and the emitter-base junction is not forward biased. Therefore, the ohmmeter should read a high resistance, as indicated on the meter. When the switch is closed, the emitter-base circuit is forward biased by the voltage across R1 and R2. Current now flows in the emitter-collector circuit which causes a lower resistance reading on the ohmmeter. A 10-to-1 resistance ratio in this test between meter readings indicates a normal gain for an audio frequency transistor.

To test an NPN transistor using this circuit, simply reverse the ohmmeter leads and carry out the procedure described earlier.

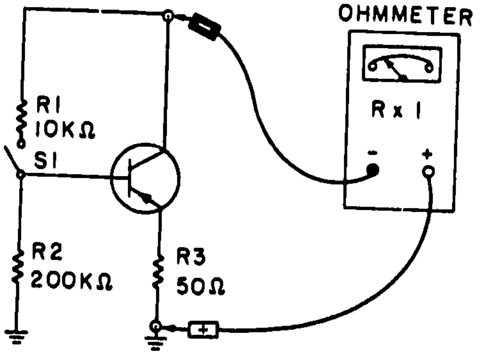


Fig 2-21. Testing a transistor's gain with an ohmmeter.

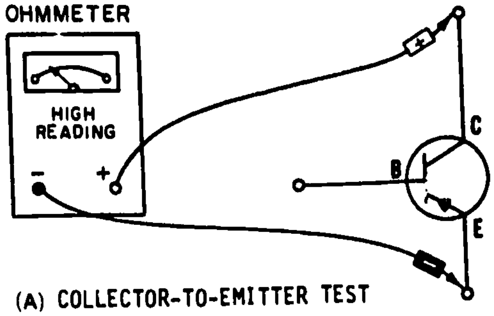
TRANSISTOR LEAKAGE TEST.-- An ohmmeter can be used to test a transistor for leakage (an undesirable flow of current) by measuring the base-emitter, base-collector, and collector-emitter forward and reverse resistances.

For simplicity, consider the transistor under test in each view of figure 2-22 as two diodes connected back to back. Therefore, each diode will have a low-forward resistance and a high-reverse resistance. By measuring their resistances with an ohmmeter as shown in the figure, you can determine if the transistor is leaking current through its junctions. When making these measurements, avoid using the R x 1 scale on the meter or a meter with a high internal battery voltage. Either of these conditions can damage a low-power transistor.

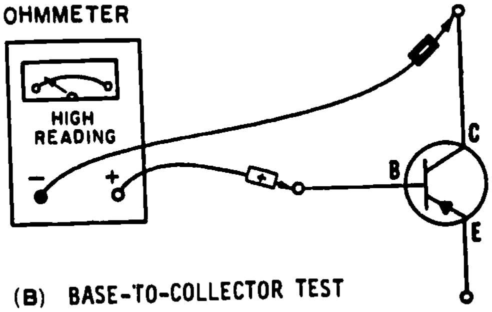
Now consider the possible transistor problems that could exist if the indicated readings in figure 2-22 are not obtained. A list of these problems is provided in the table below.

Table 2-2.

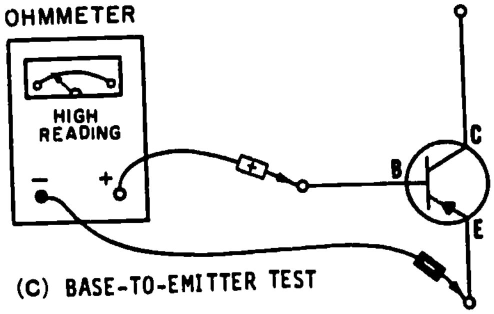
RESISTANCE READINGS		PROBLEMS
FORWARD	REVERSE	The transistor is:
	LOW (Not shorted)	LEAKING
LOW (Shorted)	LOW (Shorted)	SHORTED
HIGH	HIGH	OPEN
SAME (Nearly equal)	SAME (Nearly equal)	DEFECTIVE



(A) COLLECTOR-TO-EMITTER TEST



(B) BASE-TO-COLLECTOR TEST



(C) BASE-TO-EMITTER TEST

Note: Reversing the meter leads will give a low reading.

Fig 2-22. Testing a transistor's leakage with an ohmmeter.

By now, you should recognize that the transistor used in figure 2-22 is a PNP transistor. If you wish to test an NPN transistor for leakage, the procedure is identical to that used for testing the PNP except the readings obtained are reversed.

When testing transistors (PNP or NPN), you should remember that the actual resistance values depend on the ohmmeter scale and the battery voltage. Typical forward and reverse resistances are insignificant. The best indicator for showing whether a transistor is good or bad is the ratio of forward-to-reverse resistance. If the transistor you are testing shows a ratio of at least 30 to 1, it is probably good. Many transistors show ratios of 100 to 1 or greater.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is the greatest danger to a transistor?

____

2. What method for checking transistors is cumbersome when more than one transistor is bad in a circuit?

____

3. What safety precaution must be taken before replacing a transistor?

____

4. How is the collector lead identified on an oval-shaped transistor?

____

____

5. What are the two most important parameters used for testing a transistor?

   a. ____

   b. ____

6. When you are testing the gain of an audio-frequencey transistor with an ohmmeter, what is indicated by a 10 to 1 resistance ratio?

____

7. When you are using an ohmmeter to test a transistor for leakage, what is indicated by a low, but not shorted, reverse resistance reading?

____

Work Unit 2-7. MICROELECTRONICS

DEFINE INTEGRATED CIRCUITS.

SPECIFY THE TWO GENERAL CLASSIFICATIONS OF INTEGRATED CIRCUITS.

Up to now the various semiconductors, resistors, capacitors, etc. in our discussions have been considered as separately packaged components, called DISCRETE COMPONENTS. In this section we will introduce some of the more complex devices that contain complete circuits packaged as a single component. These devices are referred to as INTEGRATED CIRCUITS and the broad term used to describe the use of these devices to miniaturize electronic equipment is called MICROELECTRONICS.

With the advent of the transistor and the demand by the military for smaller equipment, design engineers set out to miniaturize electronic equipment. In the beginning, their efforts were frustrated because most of the other components in a circuit such as resistors, capacitors, and coils were larger than the transistor. Soon these other circuit components were miniaturized, thereby pushing ahead the development of smaller electronic equipment. Along with miniature resistors, capacitors, and other circuit elements, the production of components that were actually smaller than the space required for the interconnecting wiring and cabling became possible. The next step in the research process was to eliminate these bulky wiring components. This was accomplished with the PRINTED CIRCUIT BOARD (PCB).

A printed circuit board is a flat insulating surface upon which printed wiring and miniaturized components are connected in a predetermined design, and attached to a common base. Figure 2-23 shows a typical printed circuit board. Notice that various components are connected to the board and the printed wiring is on the reverse side. With this technique, all interconnecting wiring in a piece of equipment, except for the highest power leads and cabling, is reduced to lines of conducting material (copper, silver, aluminum, or gold) deposited directly on the surface of an insulating "circuit board." Since printed circuit boards are readily adapted as plug-in units, the elimination of terminal boards, fittings and tie points, not to mention wires, results in a substantial reduction in the overall size of electronic equipment.

After the printed circuit boards were perfected, efforts to miniaturize electronic equipment were then shifted to assembly techniques, which led to MODULAR CIRCUITRY. In this technique, printed circuit boards are stacked and connected together to form a module. This increases the packaging density of circuit components and results in a considerable reduction in the size of electronic equipment. Since the module can be designed to perform any electronic function, it is also a very versatile unit.

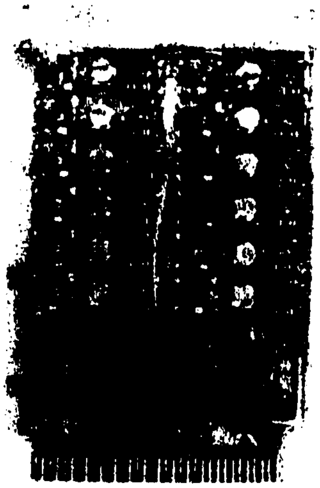
However, the drawback to this approach was that the modules required a considerable number of connections that took up too much space and increased costs. In addition tests showed the reliability was adversely affected by the increase in the number of connections.

A new technique was required to improve reliability and further increase packaging density. The solution was INTEGRATED CIRCUITS.

An integrated circuit is a device that integrates (combines) both active components (transistors, diodes, etc.) and passive components (resistors, capacitors, etc) of a complete electronic circuit in a single chip (a tiny slice or wafer of semiconductor crystal or insulator).

Integrated circuits (ICs) have almost eliminated the use of individual electronic components (resistors, capacitors, transistors, etc.) as the building blocks of electronic circuits. Instead, tiny chips have been developed whose functions are not that of a single part, but of dozens of transistors, resistors, capacitors, and other electronic elements, all interconnected to perform the task of a complex circuit. Often these comprise a number of complete conventional circuit stages, such as a multistage amplifier (in one extremely small component). These chips are frequently mounted on a plastic card called an INTEGRATED CIRCUIT BOARD (ICB), as shown in figure 2-24, which plugs into an electronic unit.

Integrated circuits have several advantages over conventionally wired circuits of discrete components. These advantages include (1) a drastic reduction in size and weight, (2) a large increase in reliability, (3) lower cost, and (4) possible improvement in circuit performance. However, integrated circuits are composed of parts so closely associated with one another that repair becomes almost impossible. In case of trouble, the entire circuit is replaced as a single component.



(A) FRONT SIDE

(B) REVERSE SIDE

Fig 2-23. A typical printed circuit board (ICB).

Fig 2-24. Typical integrated circuit board (ICB).

Basically, there are two general classifications of integrated circuits: HYBRID and MONOLITHIC. In the monolithic integrated circuit, all elements (resistors, transistors, etc.) associated with the circuit are fabricated inseparably within a continuous piece of material (called the SUBSTRATE), usually silicon. The monolithic integrated circuit is made very much like a single transistor. While one part of the crystal is being doped to form a transistor, other parts of the crystal are being acted upon to form the associated resistors and capacitor. Thus, all the elements of the complete circuit are created in the crystal by the same processes and in the same time required to make a single transistor. This produces a considerable cost savings over the same circuit made with discrete components by lowering assembly costs.

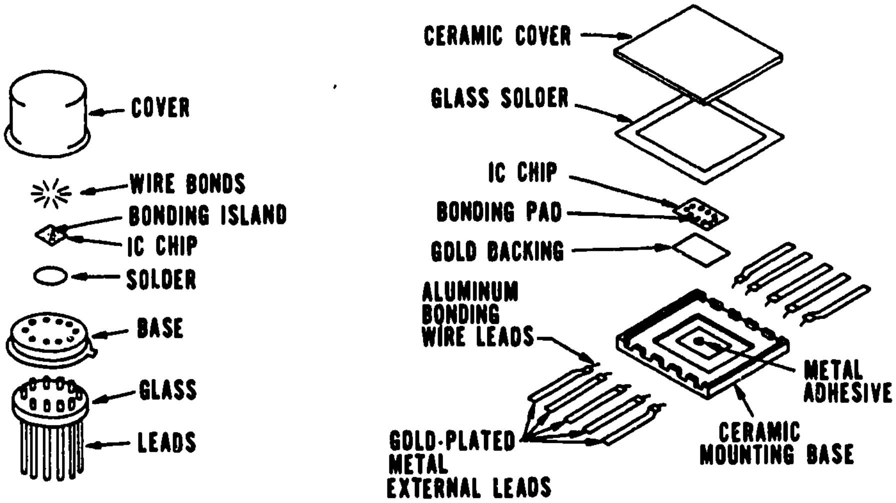


Fig 2-25. A typical integrated circuit packaging sequence.

Hybrid integrated circuits are constructed somewhat differently from the monolithic devices. The PASSIVE components (resistors, capacitors) are deposited onto a substrate (foundation) made of glass, ceramic, or other insulating material. Then the ACTIVE components (diodes, transistors) are attached to the substrate and connected to the passive circuit components on the substrate using very fine (.001 inch) wire. The term "hybrid" refers to the fact that different processes are used to form the passive and active components of the device.

Hybrid circuits are of two general types: (1) thin film and (2) thick film. "Thin" and "thick" film refer to the relative thickness of the deposited material used to form the resistors and other passive components. Thick film devices are capable of dissipating more power, but are somewhat more bulky.

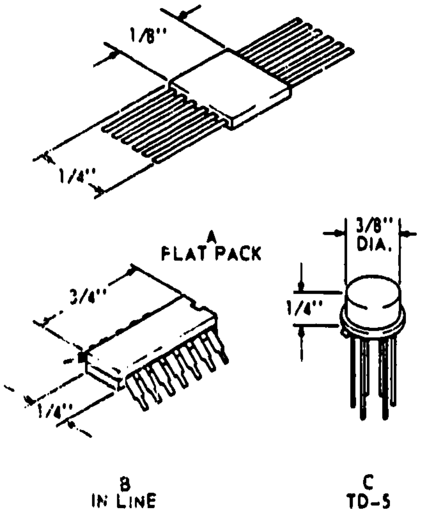


Fig 2-26. Common IC packaging styles.

Integrated circuits are being used in an ever increasing variety of applications. Small size and weight and high reliability make them ideally suited for use in airborne equipment, missile systems, computers, spacecraft, and portable equipment. They are often easily recognized because of the unusual packages that contain the integrated circuit. A typical packaging sequence is shown in figure 2-25. These tiny packages protect and help dissipate heat generated in the device. One of these packages may contain one or several stages, often having several hundred components. Some of the most common package styles are shown in figure 2-26.

The preceding information was presented to give you a brief introduction into integrated circuits. If you wish to pursue this subject further, additional information is available in your technical library.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What are integrated circuits?

2. What are the two general classifications of intergated circuits?

a.

b.

SUMMARY REVIEW

In this study unit, you learned to identify the various types of transistors, and their construction, alphanumeric descriptions, symbols and operation. You learned the precautions that must be taken when working with transistors. You learned about integrated circuits, their construction, and the advantages they offer. Lastly, you learned the function of modular circuitry.

Answers to Study Unit #2 Exercises

Work Unit 2-1.

1. Transistor
2. Amplification
3. Point-contact
4. Quality control

Work Unit 2-2.

1. Positive
2. More positive
3. Because the N material on one side of the forward-biased junction is more heavily dope than the P material.
4. The P or base section
5. 98 percent
6. Holes
7. The polarity of voltage applied to the PNP transistor is opposite of that applied to the NPN transistor.
8. I_B
9. a. Base current loop
   b. Collector current loop

Work Unit 2-3.

1. Amplifier
2. Compensation for slight variations in transistor characteristics and changes in transistor conduction due to temperature variations
3. The signals are opposite in polarity or 180 degrees out of phase with each other.
4. The polarity of the source voltage
5. Base current bias or fixed bias
6. Self-bias
7. When it is necessary to prevent amplitude
8. The voltage-divider type
9. Class A
10. Cutoff
11. a. The amount of bias
    b.The amplitude of the input signal
12. Class A

Work Unit 2-4.

1. a. Common emitter (CE)
   b. Common collector (CC)
2. Common emitter
3. Base current (IB)
4. Alpha
5. Common base
6. $V = \frac{IE}{IB}$

Work Unit 2-5.

1. a. The kind of transistor
   b. The transistor's common applications
   c. Mechanical data
2. The number of junctions in the device which in this case indicates a transistor

Work Unit 2-6.

1. Heat
2. The substitution method
3. The power must be removed from the circuit
4. By the wide space between the collector lead and the other two leads (emitter and base)
5. a. Gain
   b. Leakage
6. Normal gain
7. A leaking transistor

Work Unit 2-7.

1. An integrated circuit is a device that integrates both active and passive components of a complete electronic circuit in a single chip
2. a. Hybrid
   b. Monolithic

STUDY UNIT 3

SPECIAL DEVICES

STUDY UNIT OBJECTIVE: WITHOUT THE AID OF REFERENCES, YOU WILL IDENTIFY VARIOUS SOLID STATE DEVICES. YOU WILL ALSO IDENTIFY HOW SOLID STATE DEVICES OPERATE, HOW THEY CAN BE APPLIED AND THEIR ADVANTAGES AND DISADVANTAGES.

If you consider the sensitive nature and the various interacting properties of semiconductors, it should not be surprising to you that solid state devices can be designed for many different purposes. In fact, devices with special features are so numerous and new designs are so frequently introduced that it would be beyond the scope of this study unit to describe all of the devices in use today. Therefore, this study unit will include a variety of representative devices that are used extensively in Marine Corps equipment to give you an idea of the diversity and versatility that has been grouped into three categories: diodes, optoelectronic devices, and transistors. In this study unit each device will be described and the basic operation of each one will be discussed.

Work Unit 3-1. DIODES

EXPLAIN THE BASIC OPERATION AND THE MAJOR APPLICATIONS OF THE ZENER DIODE.

DESCRIBE THE BASIC OPERATION OF THE TUNNEL DIODE.

DESCRIBE THE BASIC OPERATION OF THE VARACTOR.

DIODES are two-terminal semiconductors of various types that are used in seemingly endless applications. The operation of normal PN-junction diodes has already been discussed, but there are a number of diodes with special properties with which you should be familiar. A discussion of all of the developments in the diode field would be impossible so some of the more commonly used special diodes have been selected for explanation. These include Zener diodes, tunnel diodes, varactors, silicon controlled rectifiers (SCR), and triacs.

When a PN junction diode is reverse biased, the majority carriers (holes in the P-material and electrons in the N-material) move away from the junction. The barrier or depletion region becomes wider, as illustrated in figure 3-1, and majority carrier current flow becomes very difficult across the high resistance of the wide depletion region. The presence of minority carriers causes a small leakage current that remains nearly constant for all reverse voltages up to a certain value. Once this value has been exceeded, there is a sudden increase in the reverse current. The voltage at which the sudden increase in current occurs is called the BREAKDOWN VOLTAGE. At breakdown, the reverse current increases very rapidly with a slight increase in the reverse voltage. Any diode can be reverse biased to the point of breakdown, but not every diode can safely dissipate the power associated with breakdown. A Zener diode is a PN junction designed to operate in the reverse-bias breakdown region.

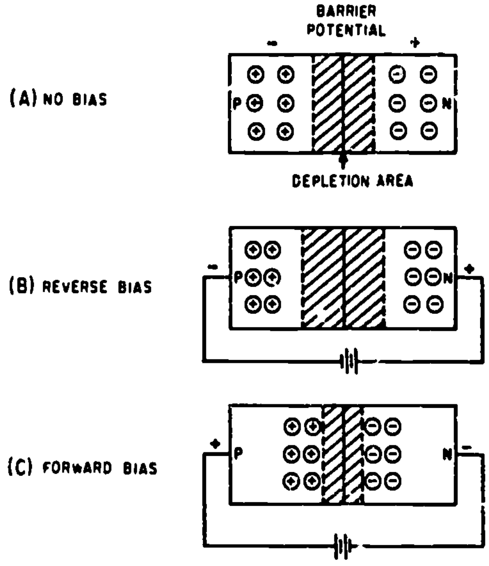


Fig 3-1. Effect of bias on the depletion region of a PN junction.

There are two distinct theories used to explain the behavior of PN junctions during breakdown: one is the ZENER EFFECT and the other is the AVALANCHE EFFECT.

The ZENER EFFECT was first proposed by Dr. Carl Zener in 1934. According to Dr. Zener's theory, electrical breakdown in solid dielectrics occurs by a process called QUANTUM-MECHANICAL TUNNELING. The Zener effect accounts for the breakdown below 5 volts; whereas, above 5 volts the breakdown is caused by the avalanche effect. Although the avalanche effect is now accepted as an explanation of diode breakdown, the term "Zener diode" is used to cover both types.

The true Zener effect in semiconductors can be described in terms of energy bands; however, only the two upper energy bands are of interest. The two upper bands, illustrated in figure 3-2, view (A), are called the conduction band and the valence band.

The CONDUCTION BAND is a band in which the energy level of the electrons is high enough that the electrons will move easily under the influence of an external field. Since current flow is the movement of electrons, the readily mobile electrons in the conduction band are capable of maintaining a current flow when an external field in the form of a voltage is applied. Therefore, solid materials which have many electrons in the conduction band are called conductors.

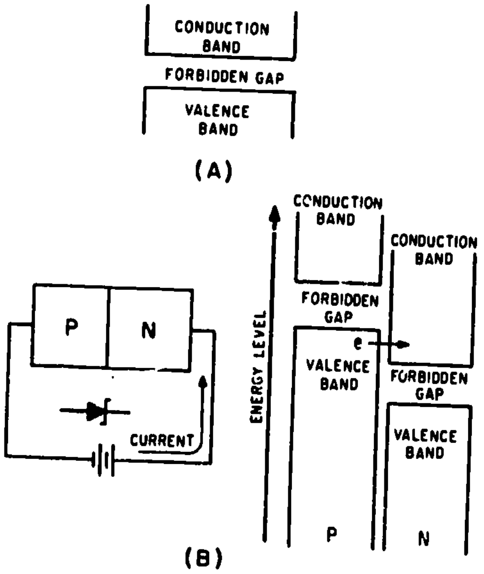


Fig 3-2. Energy diagram for Zener diode.

The VALENCE BAND is a band in which the energy level is the same as the valence electrons of the atoms. Since the electrons in these levels are attached to the atoms, the electrons are not free to move around as are the conduction band electrons. With the proper amount of energy added; however, the electrons in the valence band may be elevated to the conduction band energy level. To do this, the electrons must cross a gap that exists between the valence band energy level and the conduction band energy level. This gap is known as the FORBIDDEN ENERGY BAND or FORBIDDEN GAP. The energy difference across this gap determines whether a solid material will act as a conductor, a semiconductor, or an insulator.

A conductor is a material in which the forbidden gap is so narrow that it can be considered nonexistent. A semiconductor is a solid that contains a forbidden gap, as shown in figure 3-2, view (A). Normally, a semiconductor has no electrons at the conduction band energy level. The energy provided by room temperature heat; however, is enough energy to overcome the binding force of a few valence electrons and to elevate them to the conduction band energy level. The addition of impurities to the semiconductor material increases both the number of free electrons in the conduction band and the number of electrons in the valence band that can be elevated to the conduction band. Insulators are materials in which the forbidden gap is so large that practically no electrons can be given enough energy to cross the gap. Therefore, unless extremely large amounts of heat energy are available, these materials will not conduct electricity.

View (B) of figure 3-2 is an energy diagram of a reverse-biased Zener diode. The energy bands of the P and N materials are naturally at different levels, but reverse bias causes the valence band of the P material to overlap the energy level of the conduction band in the N material. Under this condition, the valence electrons of the P material can cross the extremely thin junction region at the overlap point without acquiring any additional energy. This action is called tunneling. When the breakdown point of the PN junction is reached, large numbers of minority carriers "tunnel" across the junction to form the current that occurs at breakdown. The tunneling phenomenon only takes place in heavily doped diodes such as Zener diodes.

The second theory of reverse breakdown effect in diodes is known as <u>AVALANCHE</u> breakdown and occurs at reverse voltages <u>beyond</u> 5 volts. This type of breakdown diode has a depletion region that is deliberately made narrower than the depletion region in the normal PN-junction diode, but thicker than that in the Zener effect diode. The thicker depletion region is achieved by decreasing the doping level from the level used in Zener effect diodes. The breakdown is a higher voltage because of the higher resistivity of the material. Controlling the doping level of the material during the manufacturing process can produce breakdown voltages ranging between about 2 and 200 volts.

The mechanism of avalanche breakdown is different from that of the Zener effect. In the depletion region of a PN junction, thermal energy is responsible for the formation of electron-hole pairs. The leakage current is caused by the movement of minority electrons, which is accelerated in the electric field, across the barrier region. As the reverse voltage across the depletion region is increased, the reverse voltage eventually reaches a critical value. Once the critical or breakdown voltage has been reached, sufficient energy is gained by the thermally released minority electrons to enable the electrons to rupture covalent bonds as they collide with lattice atoms. The released electrons are also accelerated by the electric field, resulting in the release of further electrons, and so on, in a chain or avalanche effect. This process is illustrated in figure 3-3.

For reverse voltage slightly higher than breakdown, the avalanche effect releases an almost unlimited number of carriers so that the diode essentially becomes a short circuit. The current flow in this region is limited only by an external series current-limiting resistor. Operating a diode in the breakdown region does not damage it, as long as the maximum power dissipation rating of the diode is not exceeded. Removing the reverse voltage permits all carriers to return to their normal values and velocities.

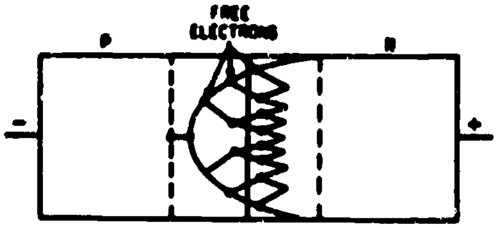


Fig. 3-3. Avalanche multiplication.

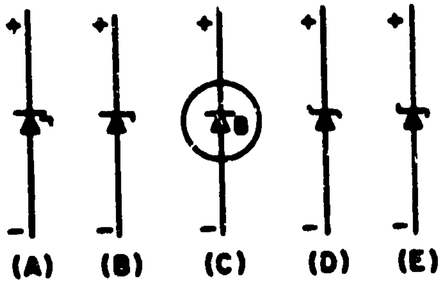


Fig 3-4. Schematic symbols for Zenor diodes.

Some of the symbols used to represent Zenor diodes are illustrated in views (A) through (E) of figure 3-4. Note that the polarity markings indicate electron flow is with the arrow symbol instead of against it as in a normal PN-junction diode. This is because breakdown diodes are operated in the reverse-bias mode which means the current flow is by minority current carriers.

Zener diodes of various sorts are used for many purposes, but their most widespread use is as voltage regulators. Once the breakdown voltage of a Zener diode is reached, the voltage across the diode then remains almost constant regardless of the supply voltage. Therefore, they hold the voltage across the load at a constant level. This characteristic makes Zener diodes ideal voltage regulators, and they are found in almost all solid state circuits in this capacity.

The Tunnel Diode.

In 1958, Leo Esaki, a Japanese scientist, discovered that if a semiconductor junction diode is heavily doped with impurities, it will have a region of negative resistance. The normal junction diode uses semiconductor materials which are lightly doped with one impurity atom for ten million semiconductor atoms. This low doping level results in a relatively wide depletion region. Conduction occurs in the normal junction diode only if the voltage applied to it is large enough to overcome the potential barrier of the junction.

In the TUNNEL DIODE, the semiconductor materials used in forming a junction are doped for the extent of one thousand impurity atoms for ten million semiconductor atoms. This heavy doping produces an extremely narrow depletion zone similar to that in the Zener diode. Also because of the heavy doping, a tunnel diode exhibits an unusual current-voltage characteristic curve as compared with that of an ordinary junction diode. The characteristic curve for a tunnel diode is illustrated in figure 3-5.

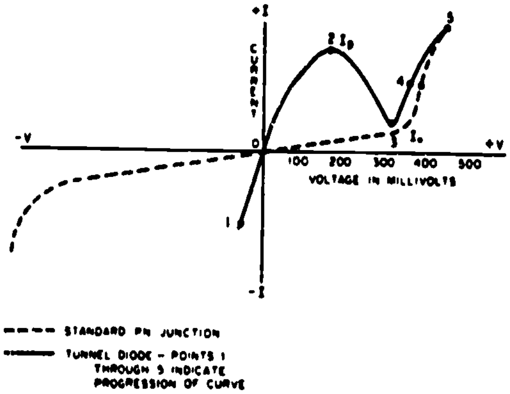


Fig 3-5. Characteristic curve of a tunnel diode compared to that of a standard PN junction.

The three most important aspects of this characteristic curve are:(1) the forward current increase to a peak (I_p) with a small applied forward bias, (2) the decreasing forward current with an increasing forward bias to a minimum valley current (I_V), and (3) the normal increasing forward current with further increases in the bias voltage. The portion of the characteristic curve between I_p and I_V is the region of negative resistance. An explanation of why a tunnel diode has a region of negative resistance is best understood by using energy levels as in the previous explanation of the Zener effect.

Simply stated the theory known as quantum-mechanical tunneling is an electron crossing a PN junction without having sufficient energy to do so otherwise. Because of the heavy doping the width of the depletion region is only one-millionth of an inch. You might think c the process simply as an arc-over between the N-and the P-side across the depletion region.

Figure 3-6 shows the equilibrium energy level diagram of a tunnel diode with no bias applied. Note in view (A) that the valence band of the P-material overlaps the conduction band of the N-material. The majority electrons and holes are at the same energy level in the equilibrium state. If there is any movement of current carriers across the depletion region due to thermal energy, the net current flow will be zero because equal numbers of current carriers flow in opposite directions. The zero net current flow is marked by an "0" on the current-voltage curve illustrated in view (B).

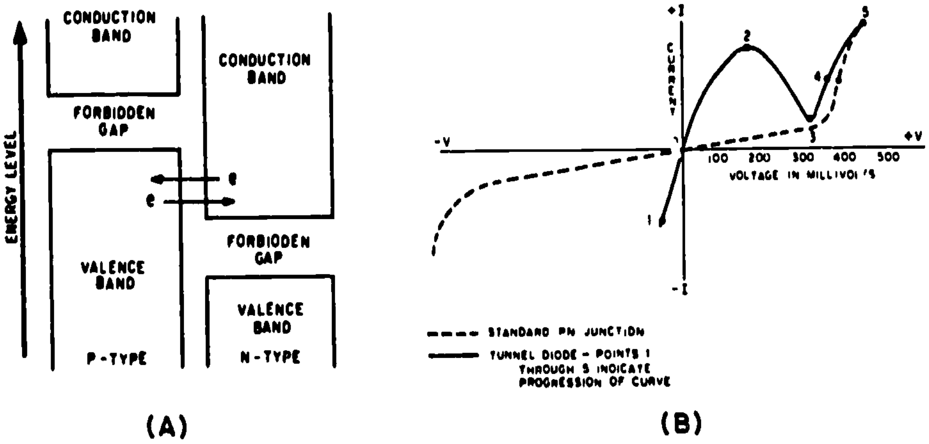


Fig 3-6. Tunnel diode energy diagram with no bias.

Figure 3-7, view (A), shows the energy diagram of a tunnel diode with a small forward bias (50) millivolts) applied. The bias causes unequal energy levels between some of the majority carriers at the energy band overlap point, but not enough of a potential difference to cause the carriers to cross the forbidden gap in the normal manner. Since the valence band of the P-material and the conduction band of the N-material still overlap, current carriers tunnel across at the overlap and cause a substantial current flow. The amount of current flow is marked by point 2 on the curve in view (B). Note in view (A) that the amount of overlap between the valence band and the conduction band decreased when forward bias was applied.

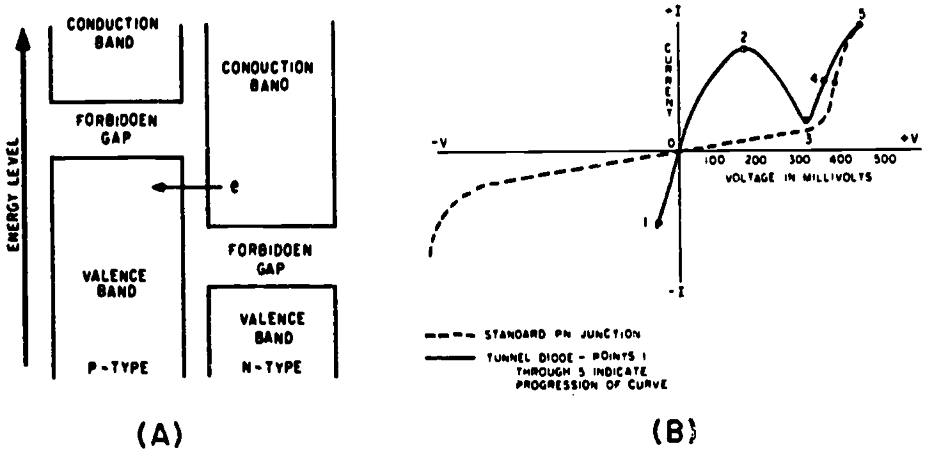


Fig 3-7. Tunnel diode energy diagram with 50 millivolts bias.

Figure 3-8, view (A), is the energy diagram of a tunnel diode in which the forward bias has been increased to 450 millivolts. As you can see, the valence band and conduction band no longer overlap at this point, and tunneling can no longer occur. The portion of the curve in view (B) from point 2 to point 3 shows the decreasing current that occurs as the bias is increased, and the area of overlap becomes smaller. As the overlap between the two energy bands becomes smaller, fewer and fewer electrons can tunnel across the junction. The portion of the curve between point 2 and point 3 in which current decreases as the voltage increases is the negative resistance region of the tunnel diode.

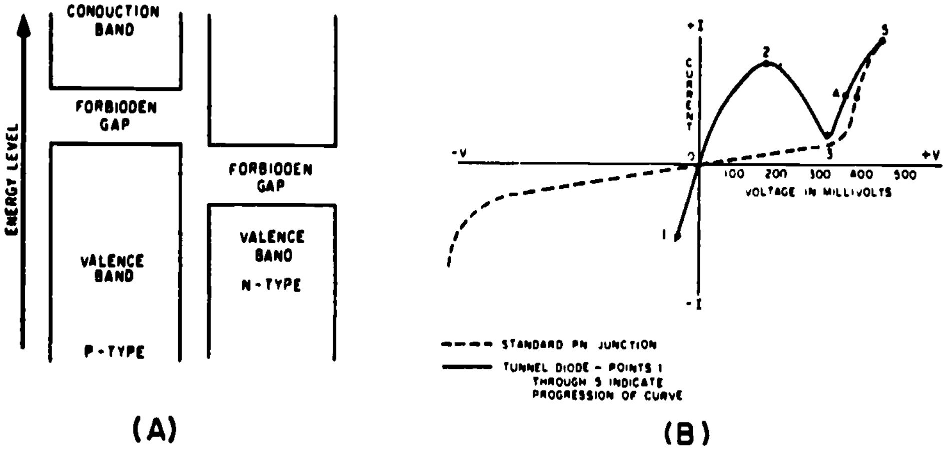


Fig 3-8. Tunnel diode energy diagram with 450 millivolts bias.

Figure 3-9, view (A), is the energy diagram of a tunnel diode in which the forward bias has been increased even further. The energy bands no longer overlap and the diode operates in the same manner as a normal PN junction, as shown by the portion of the curve in view (B) from point 3 to point 4.

The negative resistance region is the most important and most widely used characteristic of the tunnel diode. A tunnel diode biased to operate in the negative resistance region can be used as either an oscillator or an amplifier in a wide range of frequencies and applications. Very high frequency applications using the tunnel action occurs so rapidly that there is no transmit time effect and therefore no signal distortion. Tunnel diodes are also used extensively in high-speed switching circuits because of the speed of the tunneling action.

Several schematic symbols are used to indicate tunnel diodes. These symbols are illustrated in views (A) through (D) of figure 3-10.

Varactor

The VARACTOR, or varicap, as the schematic drawing in figure 3-11 suggests, is a diode that behaves like a variable capacitor, with the PN junction functioning like the dielectric and plates of a common capacitor. Understanding how the varactor operates is an important prerequisite to understanding field-effect transistors, which will be covered later in this topic.

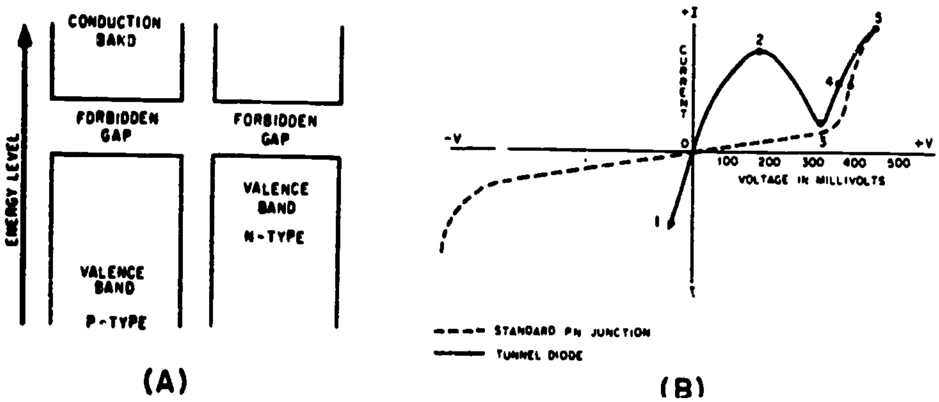


Fig 3-9. Tunnel diode energy diagram with 600 millivolts bias.

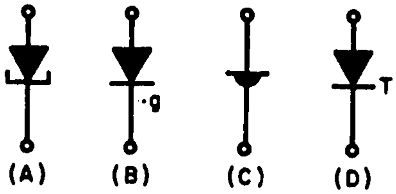


Fig 3-10. Tunnel diode schematic symbols.

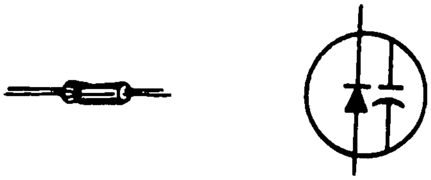


Fig 3-11. Varactor diode.

Figure 3-12 shows a PN junction. Surrounding the junction of the P and N materials is a narrow region void of both positively and negatively charged current carriers. This area is called the depletion region.

The size of the depletion region in a varactor diode is directly related to the bias. Forward biasing makes the region smaller by repelling the current carriers toward the PN junction. If the applied voltage is large enough (about .5 volt for silicon material), the negative particles will cross the junction and join with the positive particles, as shown in figure 3-13. This forward biasing causes the depletion region to decrease, producing a low resistance at the PN junction and a large current flow across it. This is the condition for a forward-biased diode. On the other hand, if reverse-bias voltage is applied to the PN junction, the size of its depletion region increases as the charged particles on both sides move away from the junction. This condition, shown in figure 3-14, produces a high resistance between the terminals and allows little current flow (only in the microampere range). This is the operating condition for the varactor diode, which is nothing more than a special PN junction.

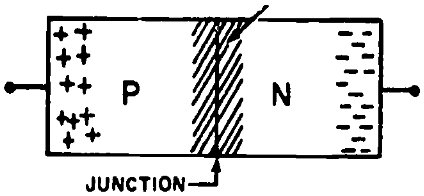


Fig 3-12. PN junction.

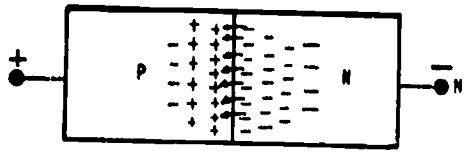


Fig 3-13. Forward-biased PN junction.

As the figure shows, the insulation gap formed by reverse biasing the varactor is comparable to the layer of dielectric material between the plates of a common capacitor. Furthermore, the formula used to calculate capacitance can be applied to both the varactor and the capacitor. In this case, the size of the insulation gap of the varactor, or depletion region, is substituted for the distance between the plates of the capacitor. By varying the reverse-bias voltage applied to the varactor, the width of the "gap" may be varied. An increase in reverse bias increases the width of the gap (d) which reduces the capacitance (C) of the PN junction. Therefore, the capacitance of the varactor is inversely proportional to the applied reverse bias.

$$C = \frac{AK}{d}$$

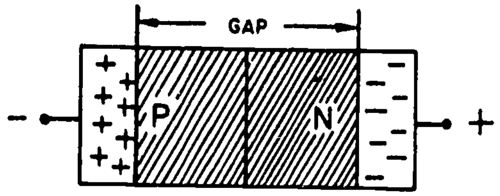
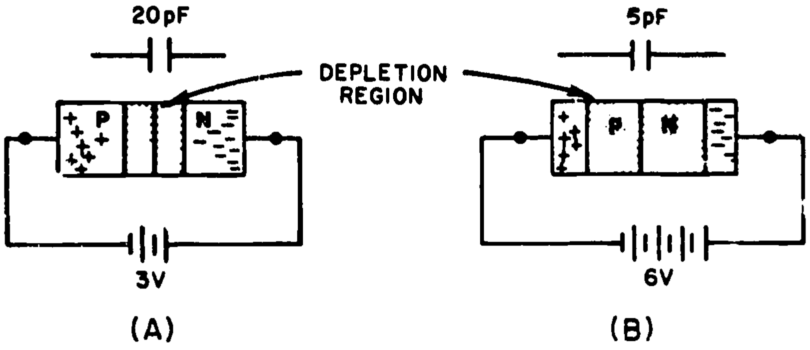


Fig 3-14. Reverse-biased PN junction.



(A)

(B)

Fig 3-15. Varactor capacitance versus bias voltage.

Where:

A = plate area

K = a constant value

d = distance between plates

The ratio of varactor capacitance to reverse-bias voltage change may be as high as 10 to 1. Figure 3-15 shows one example of the voltage-to-capacitance ratio. View (A) shows that a verse bias of 3 volts produces a capacitance of 20 picofarads in the varactor. If the reverse bias is increased to 6 volts, as shown in view (B), the depletion region widens and capacitance drops to 5 picofarads. Each 1-volt increase in bias voltage causes a 5-picofarad decrease in the capacitance of the varactor; the ratio of change is therefore 5 to 1. Of course any decrease in applied bias voltage would cause a proportionate increase in capacitance, as the depletion region narrows. Notice that the value of the capacitance is small in the picofarad range.

In general, varactors are used to replace the old style variable capacitor tuning. They are used in tuning circuits of more sophisticated communication equipment and in other circuits where variable capacitance is required. One advantage of the varactor is that it allows a d.c. voltage to be used to tune a circuit for simple remote control or automatic tuning functions. One such application of the varactor is as a variable tuning capacitor in a receiver or transmitter tank circuit like that shown in figure 3-16.

Figure 3-16 shows a d.c. voltage felt at the wiper of potentiometer R1 which can be adjusted between +V and -V. The d.c. voltage, passing through the low resistance of radio frequency choke L2, acts to reverse bias varactor diode C3. The capacitance of C3 is in series with C2, and the equivalent capacitance of C2 and C3 is in parallel with tank circuit L1-C1. Therefore, any variation in the d.c. voltage at R1 will vary both the capacitance of C3 and the resonant frequency of the tank circuit. The radio-frequency choke provides high inductive reactance at the tank frequency to prevent tank loading by R1. C2 acts to block d.c. from the tank as well as to fix the tuning range of C3.

An ohmmeter can be used to check a varactor diode in a circuit. A high reverse-bias resistance and a low forward-bias resistance with a 10 to 1 ratio in reverse-bias to forward-bias resistance is considered normal.

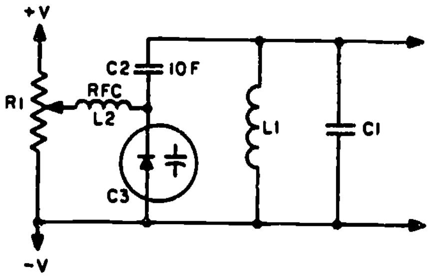


Fig 3-16. Varactor tuned.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. In a reverse-biased PN junction, which current carriers cause leakage current?

2. The action of a PN junction during breakdown can be explained by what two theories?

   a.

   b.

3. Which breakdown theory explains the action that takes place in a heavily doped PN junction with a reverse bias of less than 5 volts?

4. What is the doping level of an avalanche effect diode when compared to the doping level of a zener effect diode?

5. During avalanche effect breakdown, what limits current flow through the diode?

6. Why is electron flow with the arrow in the symbol of a zener diode instead of against the arrow as it is in a normal diode?

7. What is the main difference in construction between normal PN junction diodes and tunnel diodes?

8. What resistance property is found in tunnel diodes but not in normal diodes?

9. When compared to the ordinary diode, the tunnel diode has what type of depletion region?

10. In the tunnel diode, the tunneling current is at what level when the forbidden gap of the N-type material is at the same energy level as the empty states of the P-type material?

11. The varactor displays what useful electrical property?

12. When a PN junction is forward biased, what happens to the depletion region?

13. When the reverse bias on a varactor is increased what happens to the effective capacitance?

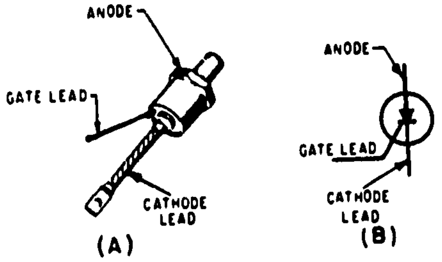
Work Unit 3-2. SILICON CONTROLLED RECTIFIER (SCR)

EXPLAIN THE BASIC OPERATION OF THE SILICON CONTROLLED RECTIFIER.

EXPLAIN THE BASIC OPERATION OF THE TRIAC.

COMPARE THE ADVANTAGES AND DISADVANTAGES OF THE SILICON CONTROLLED RECTIFIER AND TRIAC.

The SILICON CONTROLLED RECTIFIER, usually referred to as an SCR, is one of the family of semiconductors that includes transistors and diodes. A drawing of an SCR and its schematic representation is shown in views (A) and (B) of figure 3-17. Not all SCRs use the casing shown, but this is typical of most of the high-power units.



A. A high power unit.
B. The schematic symbol.

Fig 3-17. Silicon controlled rectifier.

Although it is not the same as either a diode or a transistor, the SCR combines features of both. Circuits using transistors or rectifier diodes may be greatly improved in some instances through the use of SCRs.

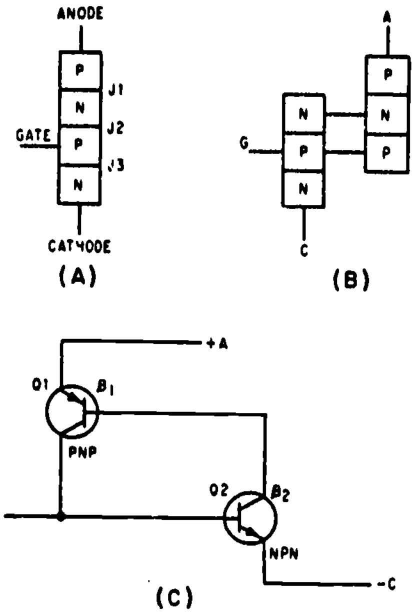
The basic purpose of the SCR is to function as a switch that can turn on or off small or large amounts of power. It performs this function with no moving parts that wear out and no points that require replacing. There can be a tremendous power gain in the SCR; in some units a very small triggering current is able to switch several hundred amperes without exceeding its rated abilities. The SCR can often replace much slower and larger mechanical switches. It even has many advantages over its more complex and larger electronic-tube equivalent, the thyratron.

The SCR is an extremely fast switch. It is difficult to cycle a mechanical switch several hundred times a minute, yet, some SCRs can be switched 25,000 times a second. It takes just microseconds (millionths of a second) to turn on or off these units. Varying the time that a switch is on as compared to the time that it is off regulates the amount of power flowing through the switch. Since most devices can operate on pulses of power (alternating current is a special form of alternating positive and negative pulses), the SCR can be used readily in controlled applications. Motor-speed controllers, inverters, remote switching units, controlled rectifiers, circuit overload protectors, latching relays, and computer logic circuits all use the SCR.

The SCR is made up of four layers of semiconductor material arranged PNPN. The construction is shown in view (A) of figure 3-18. In function, the SCR has much in common with a diode, but the theory of operation of the SCR is best explained in terms of transistors.

Consider the SCR as a transistor pair, one PNP and the other NPN, connected as shown in views (B) and (C). The anode is attached to the upper P-layer; and the gate terminal, G, goes to the P-layer of the NPN triode.

In operation, the collector of Q2 drives the base of Q1, while the collector of Q1 feeds back to the base of Q2. B1 (Beta) is the current gain of Q1, and B2 is the current gain of Q2. The gain of this positive feedback loop is their product, B1 times B2. When the product is less than one, the circuit is stable; if the product is greater than unity, the circuit is regenerative. A small negative current applied to terminal G will bias the NPN transistor into cutoff, and the loop gain is less than unity. Under these conditions, the only current that can exist between output terminals A and C is the very small cutoff collector current of the two transistors. For this reason the impedance between A and C is very high.



A. Parts of an SCR.
B. Two-transistor equivalent.
C. Two-transistor schematic.

Fig 3-18. SCR structure.

When a positive current is applied to terminal G, transistor Q2 is biased into conduction, causing its collector current to rise. Since the current gain of Q2 increases with increased collector current, a point (called the breakdown point) is reached where the loop gain equals unity and the circuit becomes regenerative. At this point, collector current of the two transistors rapidly increases to a value limited only by the external circuit. Both transistors are driven into saturation, and the impedance between A and C is very low. The positive current applied to terminal G, which served to trigger the self-regenerative action, is no longer required since the collector of PNP transistor Q1 now supplies more than enough current to drive Q2. The circuit will remain on until it is turned off by a reduction in the collector current to a value below that necessary to maintain conduction.

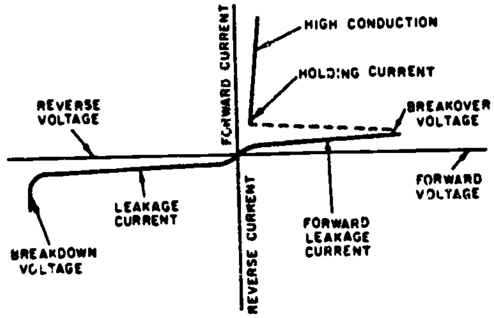


Fig 3-19. Characteristic curve for an SCR.

The characteristic curve for the SCR is shown in figure 3-19. With no gate current, the leakage current remains very small as the forward voltage from cathode to anode is increased until the breakdown point is reached. Here the center junction breaks down, the SCR begins to conduct heavily, and the drop across the SCR becomes very low.

The effect of a gate signal on the firing of an SCR is shown in figure 3-20. Breakdown of the center junction can be achieved at speeds approaching a microsecond by applying an appropriate signal to the gate lead, while holding the anode voltage constant. After breakdown, the voltage across the device is so low that the current through it from cathode to anode is essentially determined by the load it is feeding.

The important thing to remember is that a small current from gate to cathode can fire or trigger the SCR, changing it from practically an open circuit to a short circuit. The only way to change it back again (to commutate it) is to reduce the load current to a value less than the minimum forward-bias current. Gate current is required only until the anode current has completely built up to a point sufficient to sustain conduction (about 5 microseconds in resistive-load circuits). After conduction from cathode to anode begins, removing the gate current has no effect.

The basic operation of the SCR can be compared to that of the thyratron. The thyratron is an electron tube, normally gas filled, that uses a filament or a heater. The SCR and the thyratron function in a very similar manner. Figure 3-21 shows the schematic of each with the corresponding elements labeled. In both types of devices, control by the input signal is lost after they are triggered. The control grid (thyratron) and the gate (SCR) have no further effect on the magnitude of the load current after conduction begins. The load current can be interrupted by one or more of three methods (1) the load circuit must be opened by a switch, (2) the plate (anode) voltage must be reduced below the ionizing potential of the gas (thyratron). (3) the forward-bias current must be reduced below a minimum value required to sustain conduction (SCR). The input resistance of the SCR is relatively low (approximately 100 ohms) and requires a current for triggering; the input resistance of the thyratron is exceptionally high, and requires a voltage input to the grid for triggering action.

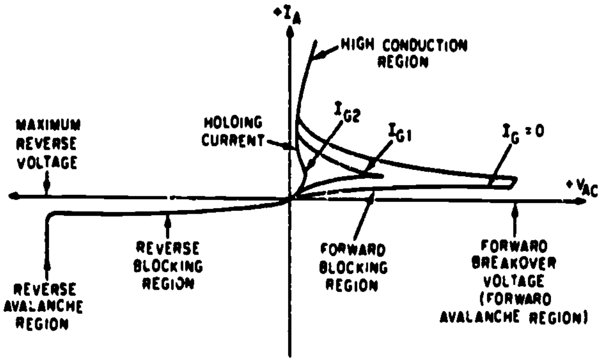
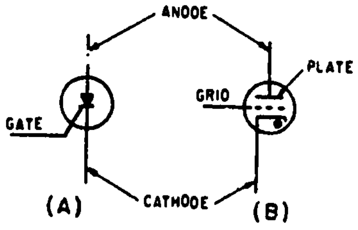


Fig 3-20. SCR characteristic curve with various gate signals.



A. SCR. B. Thyratron

Fig 3-21. Comparison of an SCR and a thyratron.

The applications of the SCR as a rectifier are many. In fact, its many applications as a rectifier give this semiconductor device its name. When alternating current is applied to a rectifier, only the positive or negative halves of the sine waves flow through. All of each positive or negative half cycle appears in the output. When an SCR is used, however, the controlled rectifier may be turned on at any time during the half cycle, thus controlling the amount of d.c. power available from zero to maximum, as shown in figure 3-22. Since the output is actually d.c. pulses, suitable filtering can be added if continuous direct current is needed. Thus any d.c. operated device can have controlled amounts of power applied to it. Notice that the SCR must be turned on at the desired time for each cycle.

When an a.c. power source is used, the SCR is turned off automatically, since the current and voltage drop to zero every half cycle. By using one SCR on positive alternations and one on negative, full-wave rectification can be accomplished, and control is obtained over the entire sine wave. The SCR serves in this application just as its name implies-as a controlled rectifier of a.c. voltage.

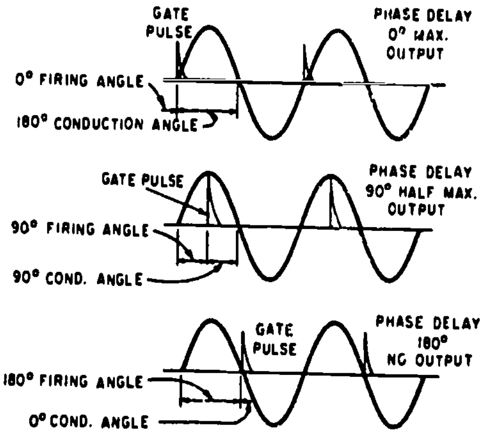


Fig 3-22. SCR gate control signals.

The TRIAC is a three-terminal devise similar in construction and operation to the SCR. The triac controls and conducts current flow during both alternations of an a.c. cycle, instead of only one. The schematic symbols for the SCR and the triac are compared in figure 3-23. Both the SCR and the triac have a gate lead. However, in the triac the lead on the same side as the gate is "main terminal 1," and the lead opposite the gate is "main terminal 2." This method of lead labeling is necessary because the triac is essentially two SCRs back to back, with a common gate and common terminals. Each terminal is, in effect, the anode of one SCR and the cathode of another, and either terminal can receive an input. In fact, the functions of a triac can be duplicated by connecting two actual SCRs as shown in figure 3-24. The result is a three-terminal device identical to the triac. The common anode-cathode connections form main terminals 1 and 2, and the common gate forms terminal 3.

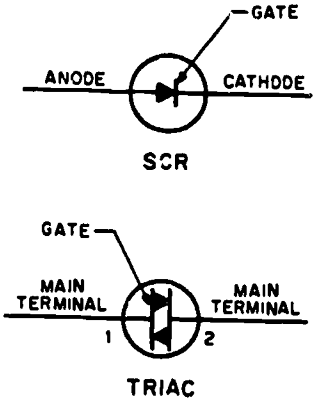


Fig 3-23. Comparison of SCR and TRIAC symbols.

The difference in current control between the SCR and the triac can be seen by comparing their operation in the basic circuit shown in figure 3-25.

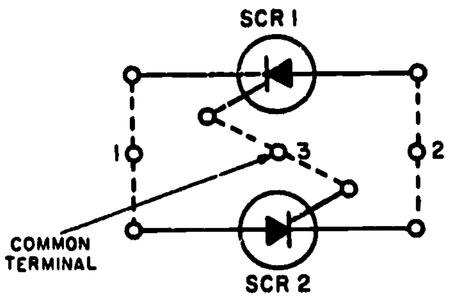


Fig 3-24. Back to back SCR equivalent circuit.

In the circuit shown in view (A), the SCR is connected in the familiar half-wave arrangement. Current will flow through the load resistor (R_L) for one alternation of each input cycle. Diode CR1 is necessary to ensure a positive trigger voltage.

In the circuit shown in the view (B), with the triac inserted in the place of the SCR, current flows through the load resistor during both alternations of the input cycle. Because either alternation will trigger the gate of the triac, CR1 is not required in the circuit. Current flowing through the load will reverse direction for half of each input cycle. To clarify this difference, a comparison of the wave forms seen at the input gate, and output points of the two devices is shown in figure 3-26.

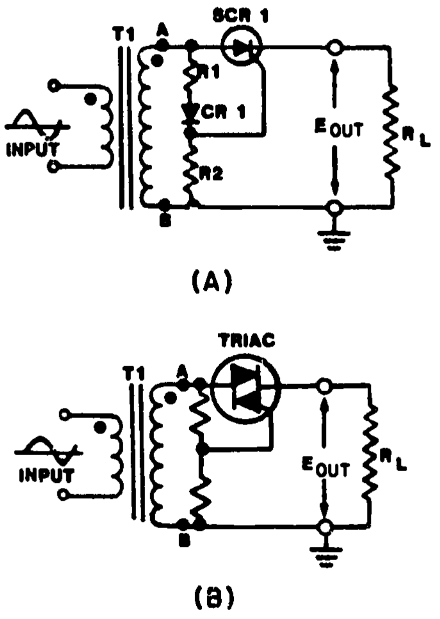


Fig 3-25. Comparison of SCR and TRIAC circuits.

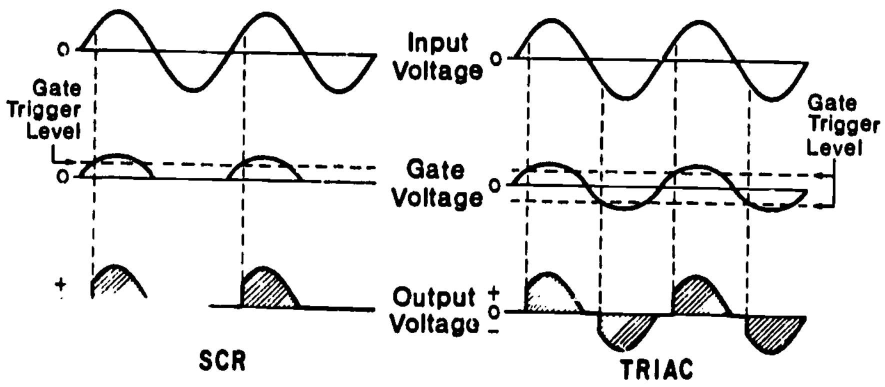


Fig 3-26. Comparison of SCR and TRIAC waveforms.

EXERCISE: Answer the following questions and check your responses against those listed at the end of the study unit.

1. The SCR is primarily used for what function?

___

2. When an SCR is forward biased, what is needed to cause it to conduct?

___

3. What is the only way to cause an SCR to stop conducting?

___

___

4. The triac is similar in operation to what device?

___

5. When used for a.c. current control, during which alternation of the a.c. cycle does the triac control current flow?

___

Work Unit 3-3. OPTOELECTRONIC DEVICES

LIST THE FIVE MOST COMMONLY USED OPTOELECTRONIC DEVICES.

EXPLAIN THE USES OF THE FIVE MOST COMMONLY USED OPTOELECTRONIC DEVICES.

OPTOELECTRONIC devices either produce light or use light in their operation. The first of these, the light emitting diode (LED), was developed to replace the fragile, short-life incandescent light bulbs used to indicate on/off conditions on panels. A LIGHT EMITTING DIODE is a diode which, when forward biased, produces visible light. The light may be red, green, or amber, depending upon the material used to make the diode.

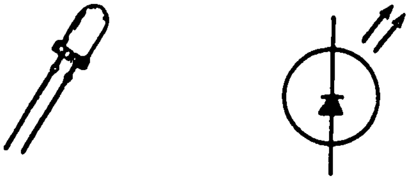
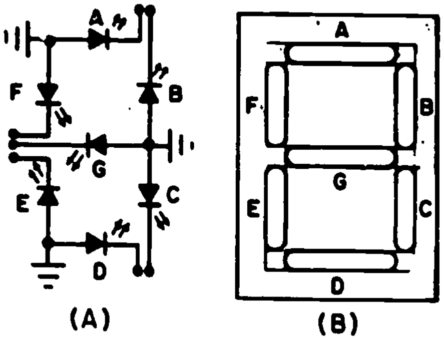


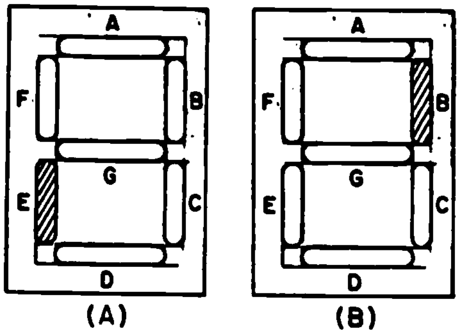
Fig 3-27. LED.



(A) (B)

Fig 3-28. Seven-segment LED display.

Figure 3-27 shows an LED and its schematic symbol. The LED is designated by a standard diode symbol with two arrows pointing away from the cathode. The arrows indicate light leaving the diode. The circuit symbols for all optoelectronic devices have arrows pointing either towards them, if they use light, or away from them, if they produce light. The LED operating voltage is small, about 1.6 volts forward bias and generally about 10 milliamperes. The life expectancy of the LED is very long, over 100,000 hours of operation.



(A) (B)

Fig 3-29. Seven segment LED display examples.

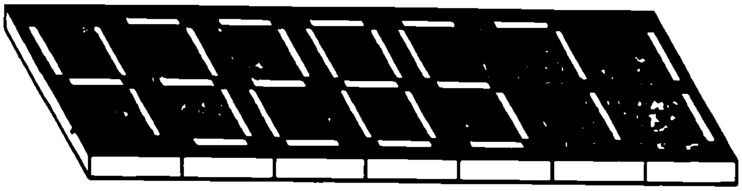


Fig 3-30. Stacked seven-segment display.

LEDs are used widely as "power on" indicators of current and as displays for pocket calculators, digital voltmeters, and frequency counters. For use in calculators and similar devices, LEDs are typically placed together in seven-segment displays, as shown in figure 3-28. This display uses seven LED segments, or bars (labeled A through G in the figure), which can be lit in different combinations to form any number from "0" through "9." The schematic, view (A), shows a common-anode display. All anodes in a display are internally connected. When a negative voltage is applied to the proper cathodes, a number is formed. For example, if negative voltage is applied to all cathodes except that of LED "E" the number "9" is produced, as shown in view (A) of figure 3-29. If the negative voltage is changed and applied to all cathodes except LED "B," the number "9" changes to "6" as shown in view (B).

Seven-segment displays are also available in common-cathode form, in which all cathodes are at the same potential. When replacing LED displays, you must ensure the replacement display is of the same type as the faulty display. Since both types look alike, you should always check the manufacturer's number.

LED seven-segment displays range from the very small, often not much larger than standard typewritten numbers, to about an inch. Several displays may be combined in a package to show a series of numbers, such as the one shown in figure 3-30.

Another special optoelectronic device in common use today is the photodiode. Unlike the LED, which produces light, the photodiode uses light to accomplish special circuit functions. Basically, the PHOTODIODE is a light-controlled variable resistor. In total darkness, it has a relatively high resistance and, therefore, conducts little current. However, when the PN junction is exposed to an external light source, internal resistance decreases and current flow increases. The photodiode is operated with reverse bias and conducts current in direct proportion to the intensity of the light source.

Figure 3-31 shows a photodiode with its schematic symbol. The arrows pointing toward the symbol indicate that light is required for operation of the device. A light source is aimed at the photodiode through a transparent "window" placed over the semiconductor chip. Switching the light source on or off changes the conduction level of the photodiode. Varying the light intensity controls the amount of conduction. Because photodiodes respond quickly to changes in light intensity, they are extremely useful in digital applications such as computer card readers, paper tape readers, and photographic light meters. They are also used in some types of optical scanning equipment.

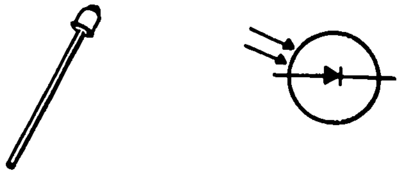


Fig 3-31. Photodiode.

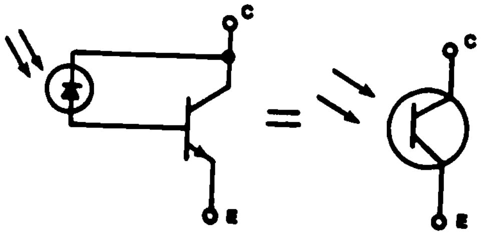


Fig 3-32. Phototransistor.

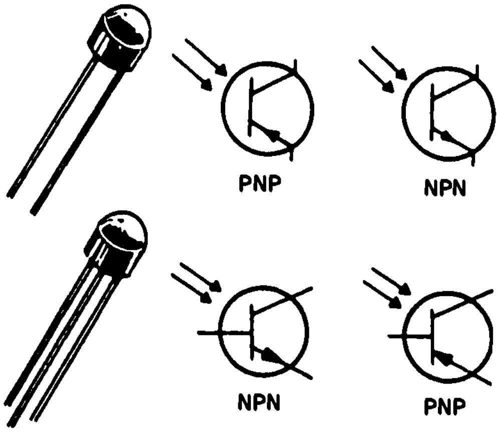


Fig 3-33. 2-Terminal and 3-terminal phototransistors.

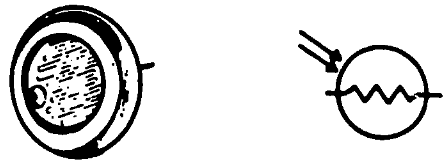


Fig 3-34. Photocell.

A second optoelectronic device, which conducts current when exposed to light, is the PHOTOTRANSISTOR. A phototransistor, however, is much more sensitive to light and produces more output current for a given light intensity than does a photodiode. Figure 3-32 shows one type of phototransistor which is made by placing a photodiode in the base circuit of an NPN transistor. Light falling on the photodiode changes the base current of the transistor, causing the collector current to be amplified. Phototransistors may also be the PNP type, with the photodiode placed in the base-collector circuit.

Figure 3-33 illustrates the schematic symbols for the various types of phototransistors. Phototransistors may be of the two-terminal type, in which the light intensity on the photodiode alone determines the amount of conduction. They may also be of the three-terminal type, which have an added base lead that allows an electrical bias to be applied to the base. The bias allows an optimum transistor conduction level, and thus compensates for ambient (normal room) light intensity.

An older device, which uses light in a way similar to the photodiode, is the photoconductive cell, or PHOTOCELL, shown with its schematic symbol in figure 3-34. Like the photodiode, the photocell is a light-controlled variable resistor. However, a typical light to dark resistance radio for a photocell is 1:1000. This means that its resistance could range from 1000 ohms in the light to 1000 kilohms in the dark, or from 2000 ohms in the light to 2000 kilohms in the dark, and so forth. Of course, other ratios are also available. Photocells are used in various types of control and timing circuits as, for example, the automatic street light controllers in most cities.

The PHOTOVOLTAIC CELL, or solar cell, is a device which converts light energy into electrical energy. An example of a solar cell and its schematic symbol are shown in figure 3-35. The symbol is similar to that of a battery. The device itself acts much like a battery when exposed to light and produces about .45 volt across its terminals, with current capacity determined by its size. As with batteries, solar cells may be connected in series or parallel to produce higher voltages and currents. The device is finding widespread application in communications satellites and solar-powered homes.

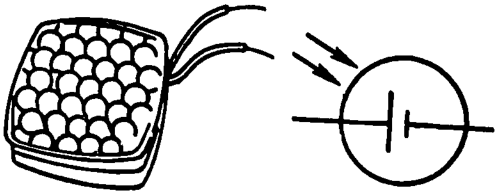


Fig 3-35. Solar cell.

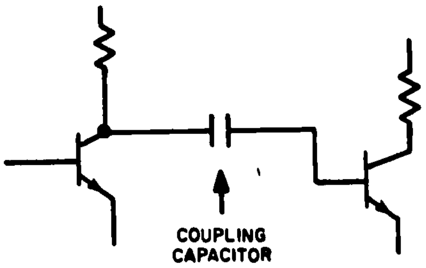


Fig 3-36. D.C blocking with a coupling capacitor.

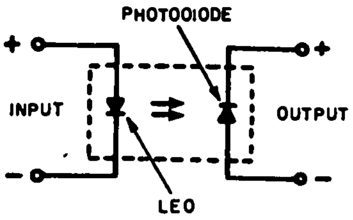


Fig 3-37. Optical coupler.

When it is necessary to block the voltage between one electronic circuit and another, and transfer the signal at the same time, an amplifier coupling capacitor is often used as shown in figure 3-36. Although this method of coupling does block d.c. between the circuits, voltage isolation is not complete. A newer method, making use of optoelectronic devices to achieve electrical isolation, is the optical coupler, shown in figure 3-37. The coupler is composed of an LED and a photodiode contained in a light-conducting medium. As the polarity signs in figure 3-37 show, the LED is forward bias, while the photodiode is reverse biased. When the input signal causes current through the LED to increase, the light produced by the LED increases. This increased light intensity causes current flow through the photodiode to increase. In this way, changes in input current produce proportional changes in the output, even though the two circuits are electrically isolated.

The optical coupler is suitable for frequencies in the low megahertz range. The photodiode type shown above can handle only small currents; however, other types of couplers, combining phototransistors with the SCR, can be used where more output is required. Optical couplers are replacing transformers in low-voltage and low-current applications. Sensitive digital circuits can utilize the coupler to control large current and voltages with low-voltage logic levels.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What type of bias is required to cause an LED to produce light?

____

2. When compared to incandescent lamps, what is the power requirement of an LED?

____

3. In a common anode, a seven-segment LED display, an individual LED will light if a negative voltage is applied to what element?

____

4. What is the resistance level of a photodiode in total darkness?

____

5. What type of bias is required for proper operation of a photodiode?

____

6. What is a typical light-to-dark resistance ratio for a photocell?

____

7. What semiconductor device produces electrical energy when exposed to light?

____

Work Unit 3-4. TRANSISTORS

DESCRIBE THE BASIC OPERATION, APPLICATIONS, AND MAJOR ADVANTAGES OF THE UNIJUNCTION TRANSISTOR.

DESCRIBE THE BASIC OPERATION, APPLICATIONS, AND MAJOR ADVANTAGES OF THE FIELD-EFFECT TRANSISTOR.

DESCRIBE THE BASIC OPERATION, APPLICATIONS AND MAJOR ADVANTAGES OF THE METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTOR.

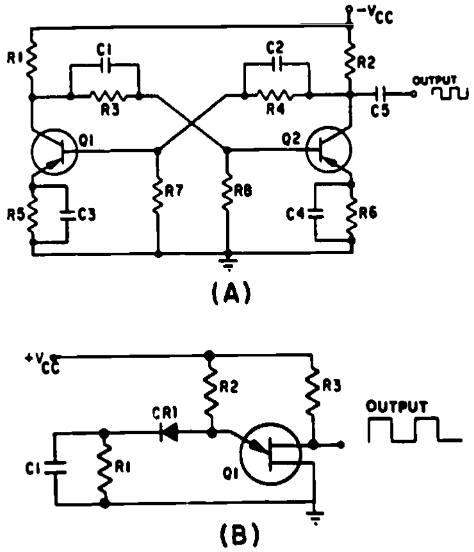
Transistors are semiconductor devices with three or more terminals. The operation of normal transistors has already been discussed, but there are several transistors with special properties that should be explained. As with diodes, a discussion of all the developments in the transistor field would be impossible. The unijunction transistor (UJT) and the field-effect transistor (FET) will be discussed because of their widespread application in Marine Corps equipment.

The UNIJUNCTION TRANSISTOR (UJT), originally called a double-based diode, is a three-terminal, solid-state device that has several advantages over conventional transistors. It is very stable over a wide range of temperatures and allows a reduction of components when used in the place of conventional transistors. A comparison is shown in figure 3-38. View (A) is a circuit using conventional transistors, and view (B) is the same circuit using the UJT. As you can see, the UJT circuit has fewer components. Reducing the number of components reduces the cost, size, and probability of failure.

The physical appearance of the UJT is identical to that of the common transistor. As shown in figure 3-39, both have three leads and the same basic shape; the tab on the case indicates the emitter on both devices. The UJT, however, has a second base instead of a collector.

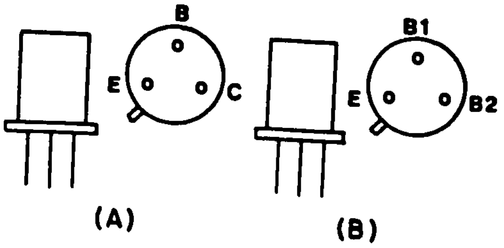
As indicated in the block diagram shown in views (A) and (B) of figure 3-40, the lead differences are even more pronounced. Unlike the transistor, the UJT has only one PN junction. The area between base 1 and base 2 acts as a resistor when the UJT is properly biased. A conventional transistor needs a certain bias level between the emitter, base, and collector for proper conduction. The same principle is true for the UJT: it needs a certain bias level between the emitter and base 1 and also between base 1 and base 2 for proper conduction.

The normal bias arrangement for the UJT is illustrated in figure 3-41, view (A). A positive 10 volts is placed on base 2 and a ground on base 1. The area between base 1 and base 2 then acts as a resistor. If a reading were taken between base 1 and base 2, the meter would indicate the full 10 volts as shown in view (B). Theoretically, if one meter lead was connected to base 1 and the other lead to some point between base 1 and base 2, the meter could read some voltage less than 10 volts. This concept is illustrated in figure 3-42, view (A). View (B) is an illustration of the voltage levels at different points between the two bases. The sequential rise in voltage is called a voltage gradient.



A. Transistor. B. UJT.

Fig 3-3B. Comparison of conventional transistor and UJT circuits.

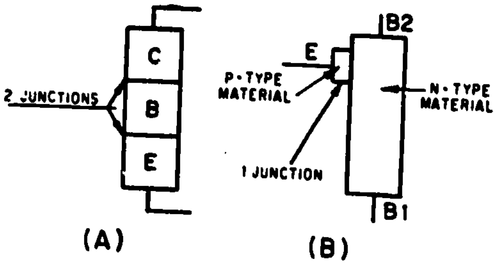


A. Transistor.

B. UJT.

Fig 3-39. Transistor and UJT.

The emitter of the UJT can be viewed as the wiper arm of a variable resistor. If the voltage level on the emitter is more positive than the voltage gradient level at the emitter-base material contact point, then the UJT is forward biased. The UJT will conduct heavily (almost a short circuit) from base 1 to the emitter. The emitter is fixed in position by the manufacturer. The level of the voltage gradient, therefore, depends upon the amount of bias voltage, as shown in figure 3-43.



A. Transistor.

B. UJT.

Fig 3-40. Transistor and UJT structure.

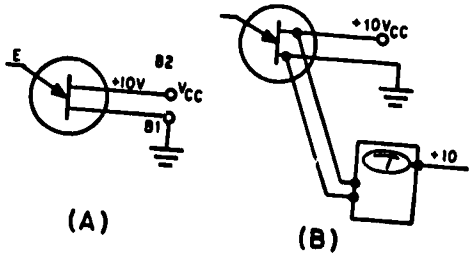


Fig 3-41. UJT biasing.

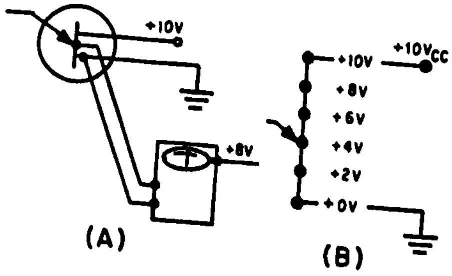


Fig 3-42. UJT voltage gradient.

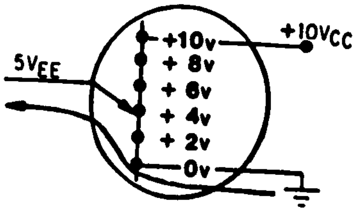


Fig 3-43. Forward bias point on UJT voltage gradient.

If the voltage level on the emitter is less positive than the voltage gradient opposite the emitter, the UJT is then reverse biased. No current will flow from base 1 to the emitter. However, a small current, called reverse current, will flow from the emitter to base 2. The reverse current is caused by the impurities used in the construction of the UJT and is in the form of minority carriers.

More than forty distinct types of UJTs are presently in use. One of the most common applications is in switching circuits. They are also used extensively in oscillators and wave shaping circuits.

Although it has brought about a revolution in the design of electronic equipment, the bipolar (PNP/NPN) transistor still has one very undesirable characteristic. The low input impedance associated with its base-emitter junction causes problems in matching impedances between interstage amplifiers.

For years, scientists searched for a solution which would combine the high input impedance of the vacuum tube with the many other advantages of the transistor. The result of this research is the FIELD-EFFECT TRANSISTOR (FET). In contrast to the bipolar transistor, which uses bias current between base and emitter to control conductivity, the FET uses voltage to control an electrostatic field within the transistor. Because the FET is voltage-controlled, much like a vacuum tube, it is sometimes called the "solid state vacuum tube."

The elements of one type of FET, the junction type (JFET), are compared with the bipolar transistor and the vacuum tube in figure 3-44. As the figure shows, the JFET is a three-element device comparable to the other two. The "gate" element of the JFET corresponds very closely in operation to the base of the transistor and the grid of the vacuum tube. The "source" and "drain" elements of the JFET correspond to the emitter and collector of the transistor and to the cathode and plate of the vacuum tube.

The construction of a JFET is shown in figure 3-45. A solid bar, made either of N-type or P-type material, forms the main body of the device. Diffused into each side of this bar are two deposits of material of the opposite type from the bar material, which form the "gate." The portion of the bar between the deposits of gate material is of a smaller cross section than the rest of the bar and forms a "channel" connecting the source and the drain. Figure 3-45 shows a bar of N-type material and a gate of P-type material. Because the material in the channel is N-type, the device is called an N-channel JFET.

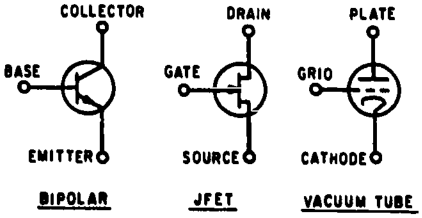


Fig 3-44. Comparison of JFET, transistor and vacuum tube symbols.

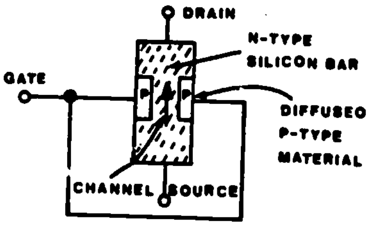


Fig 3-45. JFET structure.

In a P-channel JFET, the channel is made of P-type material and the gate of N-type material. In figure 3-46, schematic symbols for the two types of JFET are compared with those of the NPN and PNP bipolar transistor. Like the bipolar transistor types, the two types of JFET differ only in the configuration of bias voltages required and in the direction of the arrow within the symbol. Just as it does in transistor symbols, the arrow in a JFET symbol always points towards the N-type material. Thus, the symbol of the N-material JFET shows the arrow pointing towards the drain/source channel; whereas, the P-channel symbol shows the arrow pointing away from the drain/source channel, towards the gate.

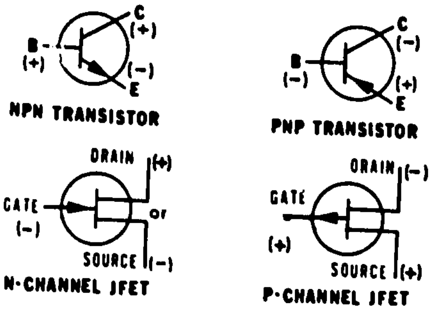


Fig 3-46. Symbols and bias voltages for transistors and JFET.

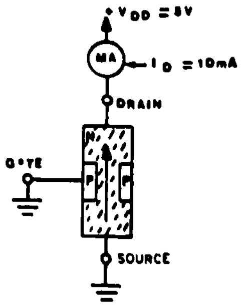


Fig 3-47. JFET operation with zero gate bias.

The key to FET operation is the effective cross-sectional area of the channel, which can be controlled by variations in the voltage applied to the gate. This is demonstrated in the figures which follow.

Figure 3-47 shows how the JFET operates in a zero gate bias condition. Five volts are applied across the JFET so that current flows through the bar from source to drain, as indicated by the arrow. The gate terminal is tied to ground. This is a zero gate bias condition. In this condition, a typical bar represents a resistance of about 500 ohms. A milliammeter, connected in series with the drain lead and d.c. power, indicates the amount of current flow. With a drain supply (V_{DD}) of 5 volts, the milliammeter gives a drain current (I_D) reading of 10 milliamperes. The voltage and current subscript letters (V_{DD}, I_D) used for an FET correspond to the elements of the FET just as they do for the elements of transistors.

In figure 3-48, a small reverse-bias voltage is applied to the gate of the JFET. A gate-source voltage (V_{GG}) of negative 1 volt applied to the P-type gate material causes the junction between the P- and N-type material to become reverse biased. Just as it did in the varactor diode, a reverse-bias condition causes a "depletion region" to form around the PN junction of the JFET. Because this region has a reduced number of current carriers, the effect of reverse biasing is to reduce the effective cross-sectional area of the "channel." This reduction in area increases the source-to-drain resistance of the device and decreases current flow.

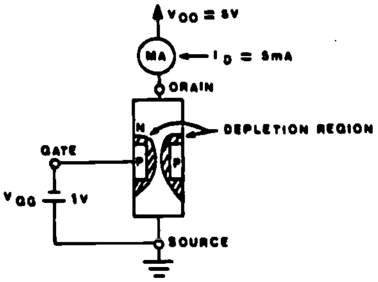


Fig 3-48. JFET with reverse bias.

The application of a large enough negative voltage to the gate will cause the depletion region to become so large that conduction of current through the bar stops altogether. The voltage required to reduce drain current (I_D) to zero is called "pinch-off" voltage and is comparable to "cuff-off" voltage in a vacuum tube. In figure 3-48, the negative 1 volt applied, although not large enough to completely stop conduction, has caused the drain current to decrease markedly (from 10 milliamperes). Calculation shows that the 1-volt gate bias has also increased the resistance of the JFET (from 500 ohms to 1 kilohm). In other words, a 1-volt change in gate voltage has doubled the resistance of the device and cut current flow in half.

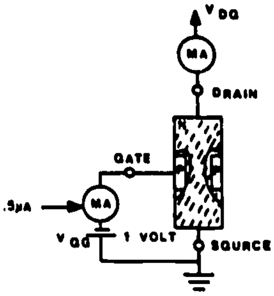


Fig 3-49. JFET input impedance.

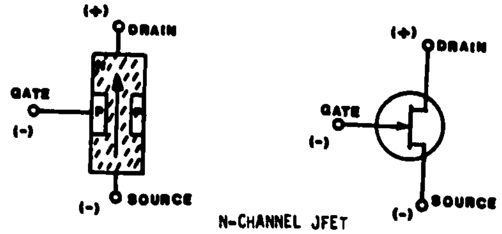
These measurements, however, show only that a JFET operates in a manner similar to a bipolar transistor, even though the two are constructed differently. As stated before, the main advantage of an FET is that its input impedance is significantly higher than that of a bipolar transistor. The higher input impedance of the JFET under reverse gate bias conditions can be seen by connecting a microammeter in series with the gate-source voltage (V_{GG}), as shown in figure 3-49.

With a V_{GG} of 1 volt, the microammeter reads .5 microamps. Applying Ohm's law (1 V ÷ .5u A) illustrates that this very small amount of current flow results in a very high input impedance (about 2 megohms). By contrast, a bipolar transistor in similar circumstances would require higher current flow (e.g., .1 to -1 mA), resulting in a much lower input impedance (about 1000 ohms or less). The higher input impedance of the JFET is possible because of the way reverse-bias gate voltage affects the cross-sectional area of the channel.

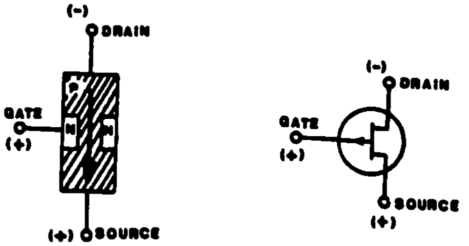
The preceding example of JFET operation uses an N-channel JFET. However, a P-channel JFET operates on identical principles. The differences between the two types are shown in figure 3-50.

Because the materials used to make the bar and the gate are reversed, source voltage potentials must also be reversed. The P-channel JFET, therefore, requires a positive gate voltage in order to be reverse biased, and current flows through it from drain to source.

Figure 3-51 shows a basic common-source amplifier circuit containing an N-channel JFET. The characteristics of this circuit include high input impedance and a high voltage gain. The function of the circuit components in this figure is very similar to those in a triode vacuum tube common-cathode amplifier circuit. C1 and C3 are the input and output coupling capacitors. R1 is the gate return resistor and functions much like the grid return resistor in a vacuum tube circuit. It prevents unwanted charge buildup on the gate by providing a discharge path for C1. R2 and C2 provide source self-bias for the JFET, which operates like cathode self-bias. R3 is the drain load resistor, which acts like the plate or collector load resistor.



N-CHANNEL JFET



P-CHANNEL JFET

Fig 3-50. JFET symbols and bias voltages.

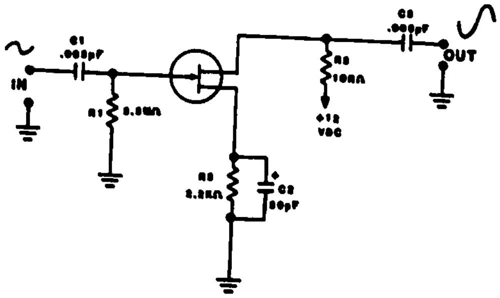
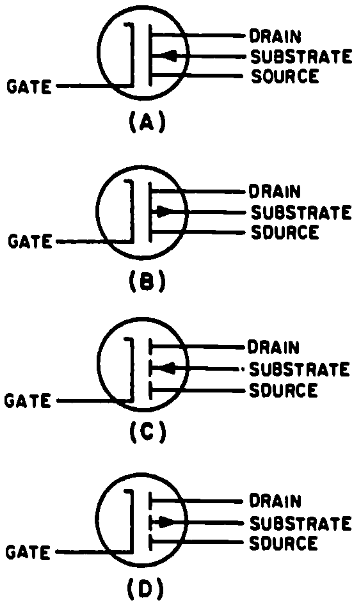


Fig 3-51. JFET common source amplifier.

The phase shift of 180 degrees between input and output signals is the same as that of common-cathode vacuum tube circuits (and common-emitter transistor circuits). The reason for the phase shift can be seen easily by observing the operation of the N-channel JFET. On the positive alternation of the input signal, the amount of reverse bias on the P-type gate materials is reduced, thus increasing the effective cross-sectional area of the channel and decreasing source-to-drain resistance. When resistance decreases, current flow through the JFET increases. This increase causes the voltage drop across R3 to increase, which in turn causes the drain voltage to decrease. On the negative alternation of the cycle, the amount of reverse bias on the gate of the JFET is increased and the action of the circuit is reversed . The result is an output signal which is an amplified 180-degree-out-of-phase version of the input signal.

A second type of field-effect transistor has been introduced in recent years that has some advantages over the JFET. This device is the metal oxide semiconductor field effect transistor (MOSFET). The MOSFET has an even higher input impedance than the JFET (10 to 100 million megohms). Therefore, the MOSFET is even less of a load on preceeding circuits. The extremely high input impedance, combined with a high gain factor, makes the MOSFET a highly efficient input device for RF/IF amplifiers and mixers and for many types of test equipment.

The MOSFET is normally constructed so that it operates in one of two basic modes: the depletion mode or the enhancement mode. The depletion mode MOSFET has a heavily doped channel and uses reverse bias on the gate to cause a depletion of current carriers in the channel. The JFET also operates in this manner. The enhancement mode MOSFET has a lightly doped channel and uses forward bias to enhance the current carriers in the channel. A MOSFET can be constructed that will operate in either mode depending upon what type of bias is applied, thus allowing a greater range of input signals.



A. N-channel, depletion, MOSFET.
B. P-channel, depletion, MOSFET.
C. N-channel, enhancement, MOSFET.
D. P-channel, enhancement, MOSFET.

Fig. 3-52. MOSFET symbols.

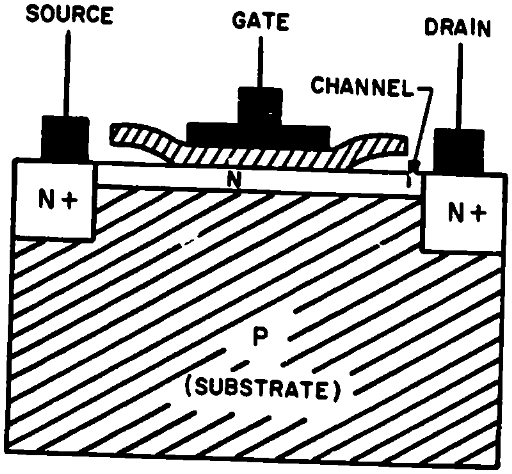
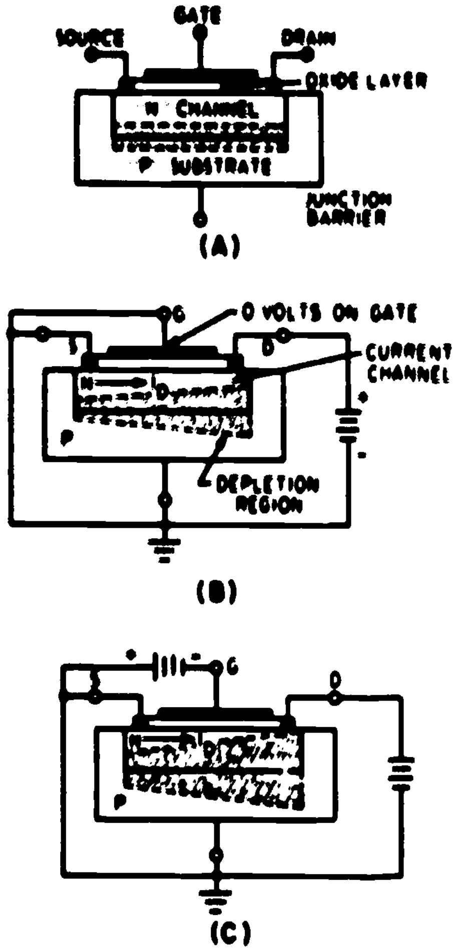


Fig 3-53. MOSFET structure.

In addition to the two basic modes of operation, the MOSFET, like the JFET, is of either the P-channel type or the N-channel type. Each type has four elements: gate, source, drain, and substrate. The schematic symbols for the four basic variations of the MOSFET are shown in views (A), (B), (C), and (D) of figure 3-52.

The construction of an N-channel MOSFET is shown in figure 3-53. Heavily doped N-type regions (indicated by the N+) are diffused into a P-type substrate or base. A channel of regular N-type material is diffused between the heavily doped N-type regions. A metal oxide insulating layer is then formed over the channel, and a metal gate layer is deposited over the insulating layer. There is no electrical connection between the gate and the rest of the device. This construction method results in the extremely high input impedance of the MOSFET. Another common name for the device, derived from the construction method, is the insulated gate field effect transistor (IGFET).

The operation of the MOSFET, or IGFET, is basically the same as the operation of the JFET. The current flow between the source and drain can be controlled by using either of two methods or by using a combination of the two methods. In one method the drain voltage controls the current when the gate potential is a zero volts. A voltage is applied to the gate in the second method. An electric field is formed by the gate voltage that affects the current flow in the channel by either depleting or enhancing the number of current carriers available. As previously stated, a reverse bias applied to the gate depletes the carriers. The polarity of the voltages required to forward or reverse bias a MOSFET depends upon whether it is of the P-channel type or the N-channel type. The effects of reverse-bias voltage on a MOSFET designed to operate in the depletion mode are illustrated in views (A), (B) and (C) of figure 3-54. The amount of reverse bias applied has a direct effect on the width of the current channel and, thus, the amount of drain current (I_D).



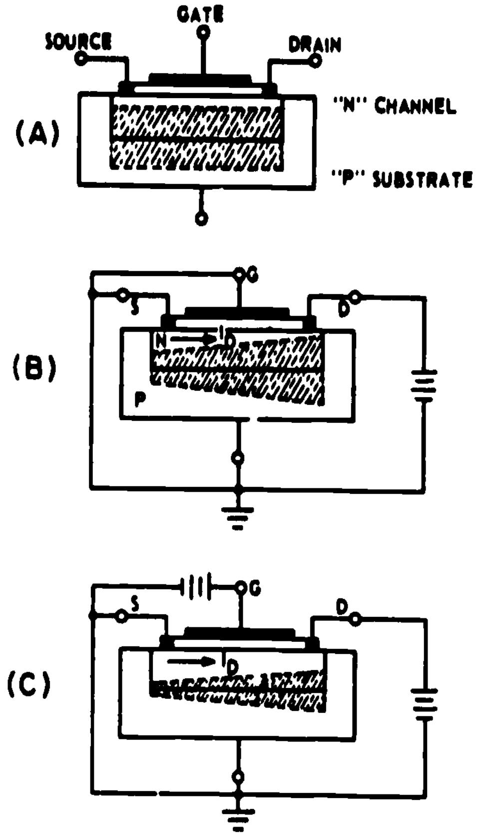
A. Substrate.
B. Source-to-drain voltage applied.
C. Reverse bias applied.

Fig 3-54. Effects of bias on N-channel depletion MOSFET.

Figure 3-56 illustrates the effect of forward bias on an enhancement mode N-channel MOSFET. In this case, a positive voltage applied to the gate increases the width of the current channel and the amount of drain current. (I_D).

Another type of MOSFET is the induced-channel type MOSFET. Unlike the MOSFETs discussed so far, the induced-channel type has no actual channel between the source and the drain. The induced channel MOSFET is constructed by making the channel of the same type material as the substrate, or the opposite of the source and the drain material. As shown in figure 3-56, the source and the drain are of P-type material, and the channel and the substrate are of N-type material.

The induced-channel MOSFET is caused to conduct from source to drain by the electric field that is created when a voltage is applied to the gate. For example, assume that a negative voltage is applied to the MOSFET in figure 3-56. The effect of the negative voltage modifies the conditions in the substrate material. As the gate builds a negative charge, free electrons are repelled, forming a depletion region. Once a certain level of depletion has occurred (determined by the composition of the substrate material), any additional gate bias attracts positive holes to the surface of the substrate. When enough holes have accumulated at the surface channel area, the channel changes from an N-type material to a P-type material, since it now has more positive carriers than negative carriers. At this point the channel is considered to be inverted, and the two P-type regions at the source and the drain are now connected by a P-type inversion layer or channel. As with the MOSFET, the gate signal determines the amount of current flow through the channel as long as the source and drain voltages remain constant. When the gate voltage is at zero, essentially no current flows since a gate voltage is required to form a channel.



A. Substrate.
B. Source-to-drain voltage applied.
C. Reverse bias applied.

Fig 3-55. Effects of bias on N-channel enhancement MOSFET.

The MOSFETs discussed up to this point have been single-gate MOSFETS. Another type of MOSFET, the dual-gate type is shown in figure 3-57. As the figure shows, the gates in a dual-gate MOSFET can be compared to the grids in a multi-grid vacuum tube. Because the substrate has been connected directly to the source terminal, the dual-gate MOSFET still has only four leads: one each for source and drain, and two for the gates. Either gate can control conduction independently, making this type of MOSFET a truly versatile device.

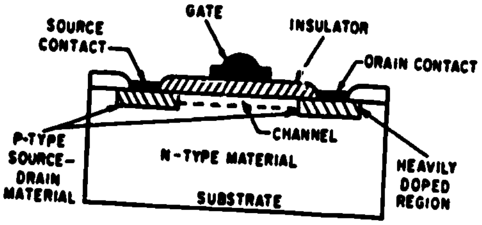


Fig 3-56. Induced channel MOSFET construction.

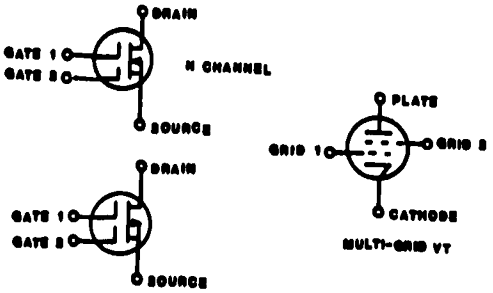


Fig 3-57. Dual gate MOSFET.

One problem with both the single- and dual-gate MOSFET is that the oxide layer between gate and channel can be destroyed very easily by ordinary static electricity. Replacement MOSFETs come packaged with their leads shorted together by a special wire loop or spring to avoid accidental damage. The rule to remember with these shorting springs is that they must not be removed until after the MOSFET has been soldered or plugged into a circuit. One such spring is shown in figure 3-58.

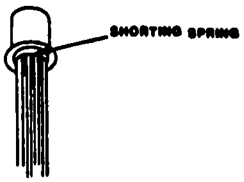


Fig 3-58. MOSFET shorting spring.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. The UJT has how many PN junctions?

2. The area between base 1 and base 2 in a UJT acts as what type of common circuit component?

3. The sequential rise in voltage between the two bases of the UJT is called what?

4. What is the normal current path for a UJT?

5. What is one of the primary advantages of the FET when compared to the bipolar transistor?

6. The FET and the vacuum tube have what in common?

7. The base cf a transistor serves a purpose similar to what element of the FET?

____

8. What are the two types of JFET?

a. _______________

b. _______________

9. The source and drain of an N-channel JFET are made of what type of material?

____

10. What is the key to FET operation?

____

11. What is the normal current path in an N-channel JFET?

____

12. Applying a reverse bias to the gate of an FET has what effect?

____

13. The input and output signals of a JFET amplifier have what phase relationship?

____

14. When compared to the JFET, what is the input impedance of the MOSFET?

____

15. What are the four elements of the MOSFET?

a. _______________

b. _______________

c. _______________

d. _______________

16. The substrate of an N-channel MOSFET is made of what material?

____

17. In a MOSFET, which element is insulated from the channel material?

____

18. What type of MOSFET can be independently controlled by two separate signals?

____

19. What is the purpose of the spring or wire around the leads of a new MOSFET?

____

SUMMARY REVIEW

This study unit introduced you to a representative selection of solid state devices that have special properties. You learned the basic operation, applications and major advantages and disadvantages of various solid state devices. It would be helpful to you to review this study unit before you proceed to the next study unit.

Answers to Study Unit #3 Exercises

Work Unit 3-1.

1. The minority carriers
2. a. Zener effect
   b. Avalanche effect
3. Zener effect
4. The doping level of an avalanche effect diode is lower
5. An external current limiting resistor
6. Because Zener diodes are operated in the reverse-bias mode
7. The amount of doping
8. Negative resistance
9. The tunnel diode has a very narrow depletion region.
10. Minimum level
11. Variable capacitance
12. The depletion region decreases.
13. Capacitance decreases

Work Unit 3-2.

1. The SCR is primarily used for switching power on or off.
2. A gate signal
3. The forward bias must be reduced below the minimum conduction level.
4. SCR
5. During both alternations

Work Unit 3-3.

1. Forward bias
2. Very low
3. The cathode
4. Very high
5. Reverse bias
6. 1:1000
7. Photovoltaic cell

Work Unit 3-4.

1. One
2. Variable resistor
3. A voltage gradient
4. From base 1 to the emitter
5. High input impedance
6. Voltage controls conduction
7. Gate
8. a. N-channel
   b. P-channel
9. N-type material
10. Effective cross-sectional area of the channel
11. From source to drain
12. Source to drain resistance decreases
13. They are 180 degrees out of phase.
14. The MOSFET has higher input impedance.
15. a. Gate
    b. Source
    c. Drain
    d. Substrate
16. P-type material
17. The gate terminal
18. The dual-gate MOSFET
19. To prevent damage from static electricity

STUDY UNIT 4

SOLID STATE POWER SUPPLIES

STUDY UNIT OBJECTIVE: WITHOUT THE AID OF REFERENCES, YOU WILL IDENTIFY THE VARIOUS COMPONENTS OF A POWER SUPPLY AND THE PURPOSE OF EACH COMPONENT. YOU WILL IDENTIFY THE VARIOUS TYPES OF RECTIFIER CIRCUITS AND FILTER CIRCUITS USED IN A POWER SUPPLY. YOU WILL IDENTIFY THE FLOW OF A.C. AND D.C. CURRENT IN A POWER SUPPLY. LASTLY, YOU WILL IDENTIFY FAULTY COMPONENTS THROUGH VISUAL CHECKS AND PROBLEMS WITHIN SPECIFIC AREAS OF A POWER SUPPLY BY USING A LOGICAL ISOLATION METHOD OF TROUBLE-SHOOTING.

In today's Marine Corps all electronic equipment require power supplies. The discovery of the silicon diode and other solid state components made possible the reduction in size and the increase in reliability of electronic equipment. This is especially important in mobile equipment where space and accessibility to spare parts are a major concern.

Work Unit 4-1. BASIC POWER SUPPLY

LIST THE VARIOUS SECTIONS OF A POWER SUPPLY.

STATE THE PURPOSE OF EACH SECTION OF A POWER SUPPLY.

DESCRIBE THE OPERATION OF THE POWER SUPPLY FROM BOTH A WHOLE UNIT STANDPOINT AND FROM A SUB-UNIT STANDPOINT.

DESCRIBE THE PURPOSE OF THE VARIOUS TYPES OF RECTIFIER CIRCUITS USED IN POWER SUPPLIES.

DESCRIBE THE PURPOSE OF THE VARIOUS TYPES OF FILTER CIRCUITS USED IN POWER SUPPLIES.

View (A) of figure 4-1 shows the block diagram of a basic power supply. Most power supplies are made up of four basic sections: a TRANSFORMER, a RECTIFIER, a FILTER, and a REGULATOR.

As illustrated in view (B) of figure 4-1, the first section is the transformer. The TRANSFORMER steps up or steps down the input line voltage and isolates the power supply from the power line. The RECTIFIER section converts the a.c. input signal to a pulsating d.c. voltage. However, you will see later that the pulsating d.c. is not desirable. For this reason, a FILTER section is used to convert pulsating d.c. to a purer, more desirable form of d.c. voltage.

The final section, the REGULATOR, does just what the name implies. It maintains the output of the power supply at a constant level in spite of large changes in load current or input line voltages.

Now that you know what each section does, let's trace an a.c. signal through the power supply. At this point you need to see how this signal is altered within each section of the power supply. Later on you will see how these changes take place. In view (B) of figure 4-1, an input signal of 115 volts a.c. is applied to the primary of the transformer. The transformer is a step-up transformer with a turns ratio of 1:3. You can calculate the output for this transformer by multiplying the input voltage by the ratio of turns in the primary to the ratio of turns in the secondary; therefore, 115 volts a.c. x 3 = 345 volts a.c. (peak-to-peak) at the output. Because each diode in the rectifier section conducts for 180 degrees of the 360-degree input, the output of the rectifier will be one-half, or approximately 173 volts of pulsating d.c. The filter section, a network of resistors, capacitors, or inductors, controls the rise and fall time of the varying signal; consequently, the signal remains at a more constant d.c. level. You will see the filter process more clearly in the discussion of the actual filter circuits. The output of the filter is a signal of 110 volts d.c., with a.c. ripple riding on the d.c. The reason for the lower voltage (average voltage) will be explained later. The regulator maintains its output at a constant 110-volt d.c. level, which is used by the electronic equipment (more commonly called the load).

114

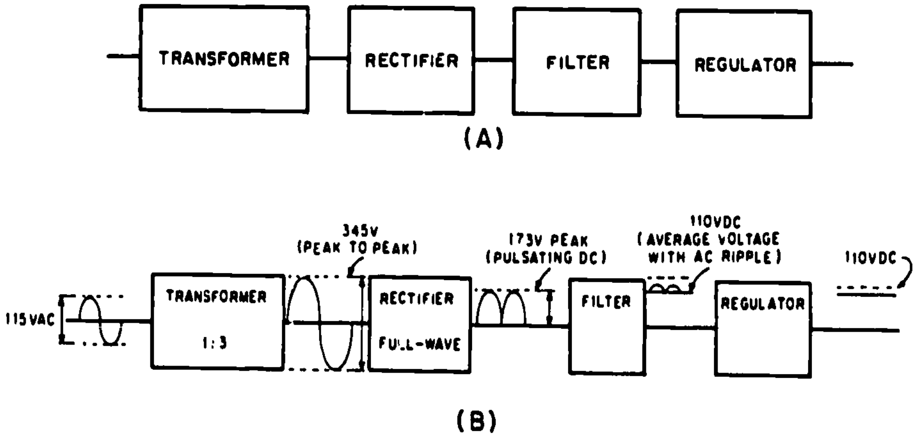
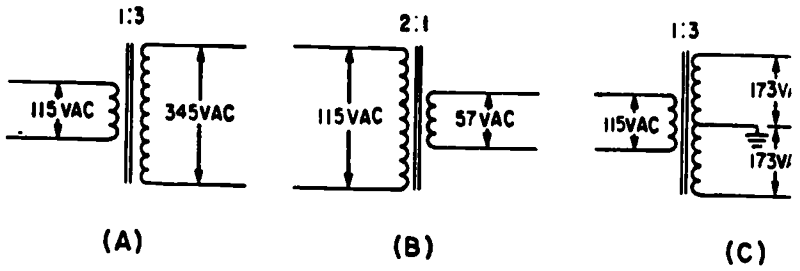


Fig 4-1. Block diagram of a basic power supply.

In some cases, a power supply may not use a transformer; therefore, the power supply would be connected directly to the source line voltage. This type of connection is used primarily because it is economical. However, unless the power supply is completely insulated, it presents a dangerous shock hazard to anyone who comes in contact with it. When a transformer is not being used, the return side of the a.c. line is connected to the metal chassis. To remove this potential shock hazard and to have the option of stepping up or stepping down the input voltage to the rectifier, a transformer must be used.

View (A) of figure 4-2 shows the schematic diagram of a STEP-UP transformer; view (B) shows a STEP-DOWN transformer; and view (C) shows a STEP-UP, CENTER-TAPPED transformer. The step-up and step-down transformers were discussed in MCI course 11.41 so only the center-tapped transformer will be mentioned in this topic. The primary purpose of the center-tapped transformer is to provide two equal voltages to the conventional full-wave rectifier.

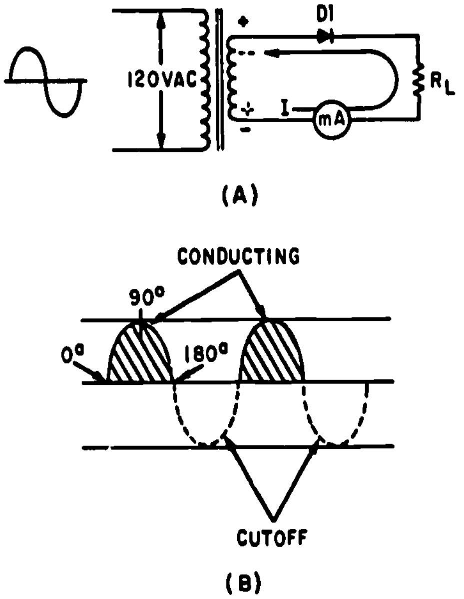
From previous discussions, you should know that rectification is the changing of an a.c. voltage to a pulsating d.c. voltage. Now let's see how the process of RECTIFICATION occurs in both a half-wave and a full-wave rectifier.



A. Step-up
B. Step-down
C. Center-tapped

Fig 4-2. Common types of transformers.

Since a silicon diode will pass current in only one direction, it is ideally suited for converting alternating current (a.c.) to direct current (d.c.). When a.c. voltage is applied to a diode, the diode conducts ONLY ON THE POSITIVE ALTERNATION OF VOLTAGE. That is, when the anode of the diode is positive in respect to the cathode. The simplest type of rectifier is the half-wave rectifier. As shown in view 'A) of figure 4-3, the half-wave rectifier uses only one diode. During the positive alternation of input voltage, the sine wave applied to the diode makes the anode positive with respect to the cathode. The diode then conducts, and current (I) flows from the negative supply lead (the secondary of the transformer), through the milliammeter, through the diode, and to the positive supply lead. As indicated by the shaded area of the output wave form in view (B), this current exists during the entire period of time that the anode is positive in respect to the cathode (in other words, for the first 180 degrees of the input sine wave).

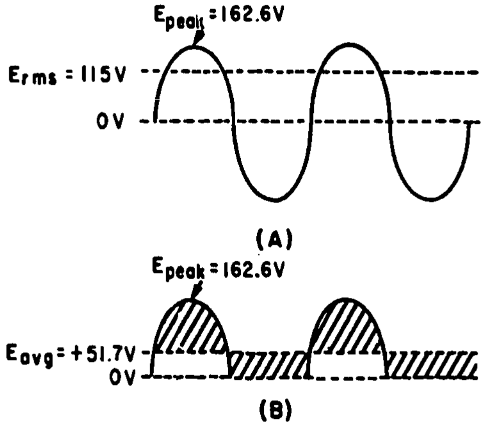


A. Half-wave rectifier
B. Output waveform

Fig 4-3. Simple half-wave rectifier.

During the negative alternation of input voltage (dotted polarity signs), the anode is driven negative and the diode cannot conduct. When conditions such as these exist, the diode is in cutoff and remains in cutoff for 18D degrees, during which time no current flows in the circuit. The circuit current, therefore, has the appearance of a series of positive pulses, as illustrated by the shaded areas on the wave form in view (B). Notice that although the current is in the form of pulses, the current always flows in the same direction. Current that flows in pulses in the same direction is called PULSATING D.C. The diode has thus RECTIFIED the a.c. input voltage.

Rms, peak, and average values. View (A) of figure 4-4 is a comparison of the rms, peak, and average values of the types of waveforms associated with the half-wave rectifier. A.C. voltages are normally specified in terms of their rms values. Thus, when a 115-volt a.c. power source is mentioned in this course, it is specifying the rms value of 115 volts a.c. In terms of peak values,



A. E_{peak} waveform
B. E_{avg} waveform

Fig 4-4. Comparison of E_{peak} to E_{avg} in a half-wave rectifier.

$$E_{rms} = E_{peak} \times .707.$$

The peak value is always higher than the rms value. In fact,

$$E_{peak} = E_{rms} \times 1.414$$

therefore, if the rms value is 115 volts a.c., then the peak value must be:

$$E_{peak} = E_{rms} \times 1.414$$

$$E_{peak} = 115 \text{ volts a.c. } \times 1.414$$

$$E_{peak}\ 162.6 \text{ volts}$$

The average value of a sine wave is 0 volts. View (B) of figure 4-4 shows how the average voltage changes when the negative portion of the sine wave is clipped off. Since the wave form swings positive but never negative (pass the "zero-volt" reference line), the average voltage is positive. The average voltage (E_{avg}) is determined by the equation:

Where: $E_{avg} = E_{peak} \times .318$

Thus: $E_{avg} = 162.6 \times .318$

$$E_{avg} = 51.7 \text{ volts}$$

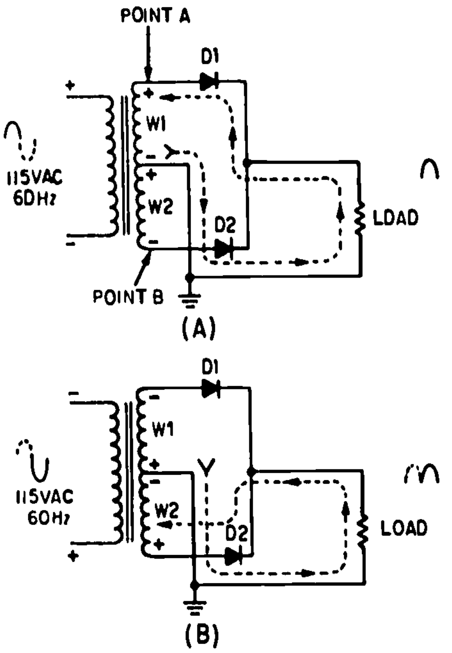
Ripple frequency. The half-wave rectifier gets its name from the fact that it conducts during only half the input cycle. Its output is a series of pulses with a frequency that is the same as the input frequency. Thus when operated from a 60-hertz line, the frequency of the pulses is 60-hertz. This is called RIPPLE FREQUENCY.

The conventional full-wave rectifier. A full-wave rectifier is a device that has two or more diodes arranged so that load current flows in the same direction during each half cycle of the a.c. supply.

A diagram of a simple full-wave rectifier is shown in figure 4-5. The transformer supplies the source voltage for two diode rectifiers, D1 and D2. This power transformer has a center-tapped, high-voltage secondary winding that is divided into two equal parts (W1 and W2). W1 provides the source voltage for D1, and W2 provides the source voltage for D2. The connections to the diodes are arranged so that the diodes conduct on alternate half cycles.

During one alternation of the secondary voltage, the polarities are as shown in view (A). The source for D2 is the voltage induced into the lower half of the secondary winding of the transformer (W2). At the specific instant of time shown in the figure, the anode voltage of D2 is negative, and D2 cannot conduct. Throughout the period of time during which the anode of D2 is negative, the anode of D1 is positive. Since the anode of D1 is positive, it conducts, causing current to flow through the load resistor in the direction shown by the arrow.

View (B) shows the next half cycle of secondary voltage. Now the polarities across W1 and W2 are reversed. During this alternation the anode of D1 is driven negative and D1 cannot conduct. For the period of time that the anode of D1 is negative, the anode of D2 is positive, permitting D2 to conduct. Notice that the anode current of D2 passes through the load resistor in the same direction as the current of D1 did. In this circuit arrangement, a pulse of load current flows during each alternation of the input cycle. Since both alternations of the input voltage cycle are used, the circuit is called a FULL-WAVE RECTIFIER.

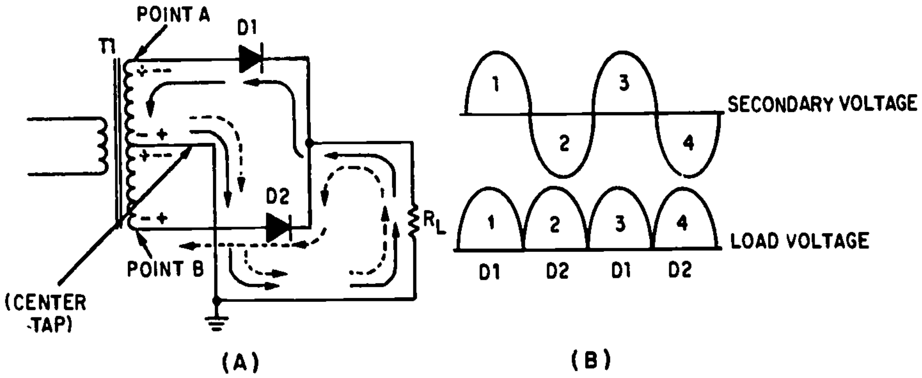


A. Positive alternation
B. Negative alternation

Fig 4-5. Full-wave rectifier.

Now that you have a basic understanding of how a full-wave rectifier works, let's cover in detail a practical full-wave rectifier and its wave forms.

A practical full-wave rectifier. A practical full-wave rectifier circuit is shown in view (A) of figure 4-6. It uses two diodes (D1 and D2) and a center-tapped transformer (T1). When the center tap is grounded, the voltages at the opposite ends of the secondary windings are 180 degrees out of phase with each other. Thus, when the voltage at point A is positive with respect to ground, the voltage at point B is negative with respect to ground. Let's examine the operation of the circuit during one comolete cycle.



A. Rectifier circuit
B. Associated waveforms

Fig 4-6. Practical full-wave rectifier.

During the first half cycle (indicated by the solid arrows), the anode of D1 is positive with respect to ground and the anode of D2 is negative. As shown, current flows from ground (center tap), up through the load resistor (R_L), through diode D1 to point A. In the transformer, current flows from point A, through the upper winding, and back to ground (center tap). When D1 conducts, it acts like a closed switch so that the positive half cycle is felt across the load (R_L).

During the second half cycle (indicated by the dotted lines), the polarity of the applied voltage has reversed. Now the anode of D2 is positive with respect to ground and the anode of D1 is negative. Now only D2 can conduct. Current now flows, as shown, from ground (center tap), up through the load resistor (R_L), through diode D2 to point B of T1. In the transformer, current flows from point B up through the lower windings and back to ground (center tap). Notice that the current flows across the load resistor (R_L) in the SAME DIRECTION for both halves of the input cycle.

View (B) represents the output wave form from the full-wave rectifier. The wave form consists of two pulses of current (or voltage) for each cycle of input voltage. The ripple frequency at the output of the full-wave rectifier is therefore TWICE THE LINE FREQUENCY.

The higher frequency at the output of a full-wave rectifier offers a distinct advantage: Because of the higher ripple frequency, the output is closely approximate to pure d.c. The higher frequency also makes filtering much easier than it is for the output of the half-wave rectifier.

In terms of peak value, the average value of current and voltage at the output of the full-wave rectifier is twice as great as that at the output of the half-wave rectifier. The relationship between the peak value and the average value is illustrated in figure 4-7. Since the output wave form is essentially a sine wave with both alternations at the same polarity, the average current or voltage is 63.7 percent (or 0.637) of the peak current or voltage.

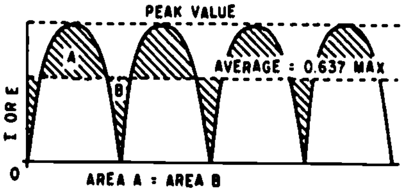


Fig 4-7. Peak and average values for a full-wave rectifier.

As an equation:

Where:

E_{max} = The peak value of the load voltage pulse

$E_{avg} = 0.637 \times E_{max}$ (the average load voltage)

I_{max} = The peak value of the load current pulse

$I_{avg} = 0.637 \times I_{max}$ (the average load current)

Example: The total voltage across the high-voltage secondary of a transformer used to supply a full-wave rectifier is 300 volts. Find the average load voltage (ignore the drop across the diode).

Solution: Since the total secondary voltage (E_s) is 300 volts, each diode is supplied one-half of this value, or 150 volts. Because the secondary voltage is an rms value, the peak load voltage is:

$$E_{max} = 1.414 \times E_s$$

$$E_{max} = 1.414 \times 150$$

$$E_{max} = 212 \text{ volts}$$

The average load voltage is:

$$E_{avg} = 0.637 \times E_{max}$$

$$E_{avg} = 0.637 \times 212$$

$$E_{avg} = 135 \text{ volts}$$

Note: If you have problems with this equation, review the portion of FM 11-61 (Communications-Electronics Fundamentals: Basic Principles of Alternating Current) that pertains to this subject.

As you may recall from your past studies in electricity, every circuit has advantages and disadvantages. The full-wave rectifier is no exception. In studying the full-wave rectifier, you may have found that by doubling the output frequency, the average voltage has doubled, and the resulting signal is much easier to filter because of the high ripple frequency. The only disadvantage is that the peak voltage in the full-wave rectifier is only half the peak voltage in the half-wave rectifier. This is because the secondary of the power transformer in the full-wave rectifier is center tapped; therefore, only half the source voltage goes to each diode.

Fortunately, there is a rectifier which produces the same peak voltage as a half-wave rectifier and the same ripple frequency as a full-wave rectifier. This circuit, known as the BRIDGE RECTIFIER, will be the subject of our next discussion.

The bridge rectifier. When four diodes are connected as shown in figure 4-8, the circuit is called a BRIDGE RECTIFIER. The input to the circuit is applied to the diagonally opposite corners of the network, and the output is taken from the remaining two corners.

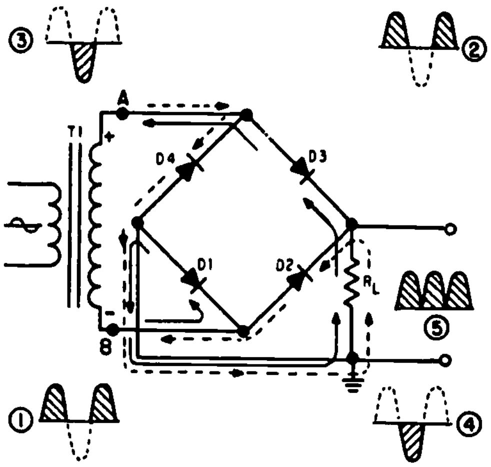
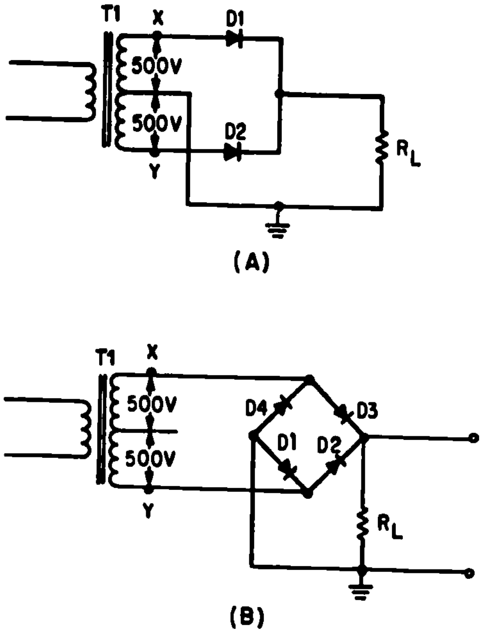


Fig 4-8. Bridge rectifier.

One complete cycle of operation will be discussed to help you understand how this circuit works. Let us assume the transformer is working properly and there is a positive potential at point A and a negative potential at point B. The positive potential at point A will forward bias D3 and reverse bias D4. The negative potential at point B will forward bias D1 and reverse bias D2. At this time D3 and D1 are forward biased and will allow current flow to pass through them; D4 and D2 are reverse biased and will block current flow. The path for current flow is from point B through D1, up through R_L, through D3, through the secondary of the transformer back to point B. This path is indicated by the solid arrows. Wave forms (1) and (2) can be observed across D1 and D3.

One-half cycle later the polarity across the secondary of the transformer reverses, forward biasing D2 and D4 and reverse biasing D1 and D3. Current flow will now be from point A through D4, up through R_L, through D2, through the secondary of T1, and back to point A. This path is indicated by the broken arrows. Waveforms (3) and (4) can be observed across D2 and D4. You should have noted that the current flow through R_L is always in the same direction. In flowing through R_L this current develops a voltage corresponding to that shown in waveforms (5). Since current flows through the load (R_L) during both half cycles of the applied voltage, this bridge rectifier is a full-wave rectifier.

One advantage of a bridge rectifier over a conventional full-wave rectifier is that with a given transformer the bridge rectifier produces a voltage output that is nearly twice that of the conventional full-wave circuit. This may be shown by assigning values to some of the components shown in views (A) and (B) of figure 4-9. Assume that the same transformer is used in both circuits. The peak voltage developed between points X and Y is 1000 volts in both circuits. In the conventional full-wave circuit shown in view (A), the peak voltage from the center tap to either X or Y is 500 volts. Since only one diode can conduct at any instant, the maximum voltage that can be rectified at any instant is 500 volts. Therefore, the maximum voltage that appears across the load resistor is nearly - but never exceeds - 500 volts, as a result of the small voltage drop across the diode. In the bridge rectifier shown in view (B), the maximum voltage that can be rectified is the full secondary voltage, which is 1000 volts. Therefore, the peak output voltage across the load resistor is nearly 1000 volts. With both circuits using the same transformer, the bridge rectifier circuit produces a higher output voltage than the conventional full-wave rectifier circuit.



A. Conventional full-wave rectifier
B. Full-wave bridge rectifier

Fig 4-9. Comparison of a conventional and bridge full-wave rectifier.

FILTERS

While the output of a rectifier is a pulsating d.c., most electronic circuits require a substantially pure d.c. for proper operation. This type of output is provided by SINGLE or multisection filter circuits placed between the output of the rectifier and the load.

There are four basic types of filter circuits:

Simple capacitor filter

LC choke-input filter

LC capacitor-input filter (pi-type)

RC capacitor-input filter (pi-type)

The function of each of these filters will be covered in detail.

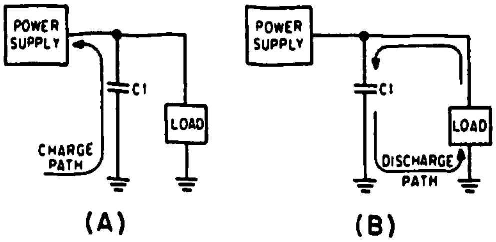
Filtering is accomplished by the use of capacitors, inductors, and/or resistors in various combinations. Inductors are used as series impedance to oppose the flow of alternating (pulsating d.c.) current. Capacitors are used as shunt elements to bypass the alternating components of the signal around the load (to ground). Resistors are used in place of inductors in low current applications.

Let's briefly review the properties of a capacitor. First, a capacitor opposes any change in voltage. The opposition to a change is voltage is called capacitive reactance (X_C) and is measured in ohms. The capacitive reactance is determined by the frequency (f) of the applied voltage and the capacitance (C) of the capacitor.

$$X_C = \frac{1}{2\pi fC} \text{ or } \frac{.159}{fC}$$

From this formula, you can see that if frequency or capacitance is increased, the X_C decreases. Since filter capacitors are placed in parallel with the load, a low X_C will provide better filtering than a high X_C. For this to be accomplished, a better shunting effect of the a.c. around the load is provided, as shown in figure 4-10.

To obtain a steady d. c. output, the capacitor must charge almost instantaneously to the value of applied voltage. Once charged, the capacitor must retain the charge as long as possible. The capacitor must have a short charge time constant (view A). This can be accomplished by keeping the internal resistance of the power supply as small as possible (fast charge time) and the resistance of the load as large as possible (for a slow discharge time as illustrated in view B).



A. Fast charge
B. Slow discharge

Fig 4-10. Capacitor filter.

From your earlier studies in basic electricity, you may remember that one time constant is defined as the time it takes a capacitor to charge to 63.2 percent of the applied voltage or to discharge to 36.8 percent of its total charge. This action can be expressed by the following equation:

$$t = RC$$

Where:

R represents the resistance to the charge or discharge path

And:

C represents the capacitance of the capacitor.

You should also recall that a capacitor is considered fully charged after five RC time constants. Refer to figure 4-11. You can see that a steady d.c. output voltage is obtained when the capacitor charges rapidly and discharges as slowly as possible.

In filter circuits, the capacitor is the common element to both the charge and the discharge paths. Therefore, to obtain the longest possible discharge time, you want the capacitor to be as large as possible. Another way to look at it is: The capacitor acts as a short circuit around the load (as far as the a.c. component is concerned), and since the larger the value of the capacitor (C), the smaller the opposition (X_C) or resistance to a.c.

$$X_C = \frac{1}{2 \pi f C}$$

Now let's look at inductors and their application in filter circuits. Remember, AN INDUCTOR OPPOSES ANY CHANGE IN CURRENT. In case you have forgotten, a change in current through an inductor produces a changing electromagnetic field. The changing field, in turn, cuts the windings of the wire in the inductor and thereby produces a counterelectromotive force (cemf). It is the cemf that opposes the change in circuit current. Opposition to a change in current at a given frequency is called inductive reactance (X_L) and is measured in ohms. The inductive reactance (X_L) of an inductor is determined by the applied frequency and the inductance of the inductor.

Mathematically,

$$X_L = 2 \pi f L$$

If frequency or inductance is increased, the X_L increases. Since inductors are placed in series with the load (as shown in figure 4-12), the larger the X_L, the larger the a.c. voltage developed across the load.

Now refer to figure 4-13. When the current starts to flow through the coil, an expanding magnetic field builds up around the inductor. This magnetic field around the coil develops the cemf that opposes the change in current. When the rectifier current decreases, as shown in figure 4-14, the magnetic field collapses and again cuts the turns (windings) of wire, thus inducing current into the coil. This additional current merges with the rectifier current and attempts to keep it at its original level.

Now that you have read how the components in a filter circuit react to current flow from the rectifier, the different types of filter circuits in use today will be discussed.

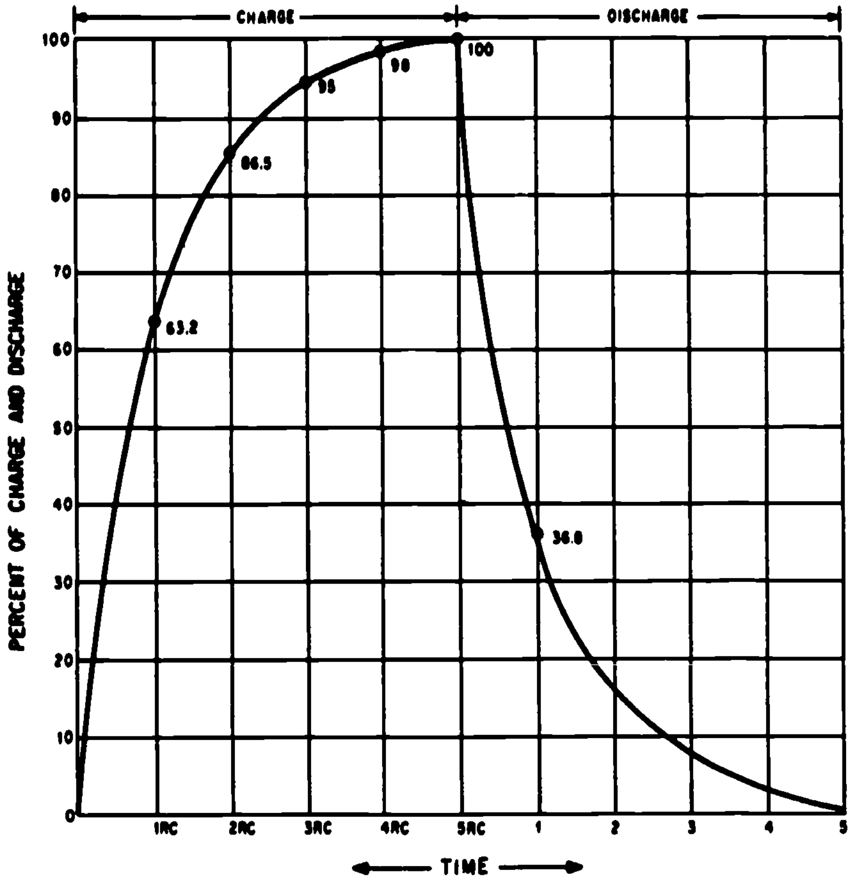


Fig 4-11. RC time constant.

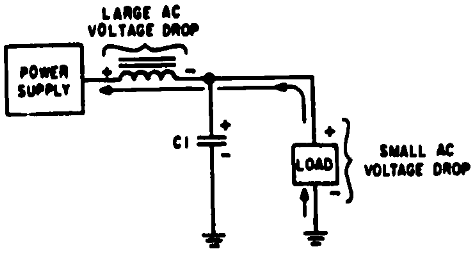


Fig 4-12. Voltage drops in an inductive filter.

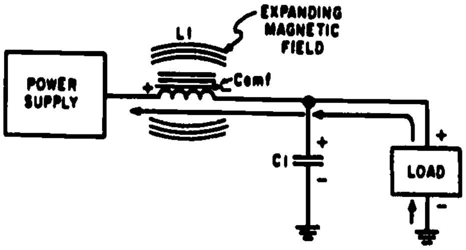


Fig 4-13. Inductive filter (expanding field).

The capacitor filter. The simple capacitor filter is the most basic type of power supply filter. The application of the simple capacitor filter is very limited. It is sometimes used on extremely high-voltage, low-current power supplies for cathode-ray and similar electron tubes which require very little load current from the supply. The capacitor filter is also used where the power-supply ripple frequency is not critical; this frequency can be relatively high. The capacitor (C1) shown in figure 4-15 is a simple filter connected across the output of the rectifier in parallel with the load.

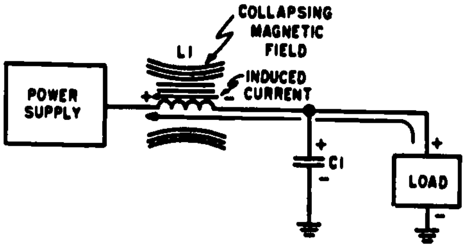


Fig 4-14. Inductive filter (collapsing field).

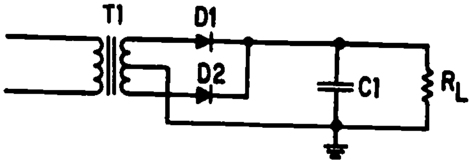
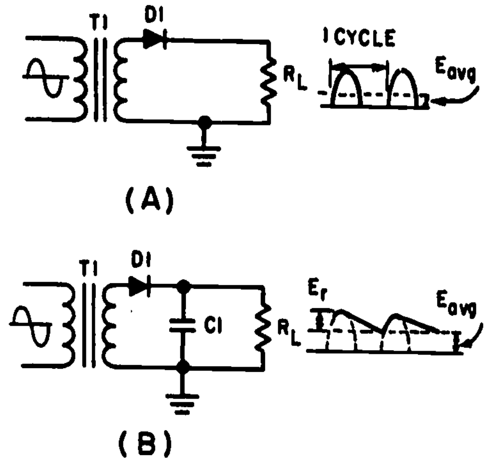


Fig 4-15. Full-wave rectifier with a capacitor filter.

When this filter is used, the RC charge time of the filter capacitor (C1) must be short and the RC discharge time must be long to eliminate ripple action. In other words, the capacitor must charge up fast, preferably with no discharge at all. Better filtering also results when the input frequency is high; therefore, the full-wave rectifier output is easier to filter than that of the half-wave rectifier because of its higher frequency.

For you to have a better understanding of the effect that filtering has on E_{avg}, a comparison of a rectifier circuit with a filter and one without a filter is illustrated in views (A) and (B) of figure 4-16. The output waveforms in figure 4-16 represent the unfiltered and filtered outputs of the half-wave rectifier circuit. Current pulses flow through the load resistance (R_L) each time a diode conducts. The dashed line indicates the average value of output voltage. For the half-wave rectifier, E_{avg} is less than half (or approximately 0.318) of the peak output voltage. This value is still much less than that of the applied voltage. With no capacitor connected across the output of the rectifier circuit, the waveform in view (A) has a large pulsating component (ripple) compared with the average or d.c. component. When a capacitor is connected across the output (view B), the average value of output voltage (E_{avg}) is increased due to the filtering action of capacitor C1.



A. Unfiltered
B. Filtered

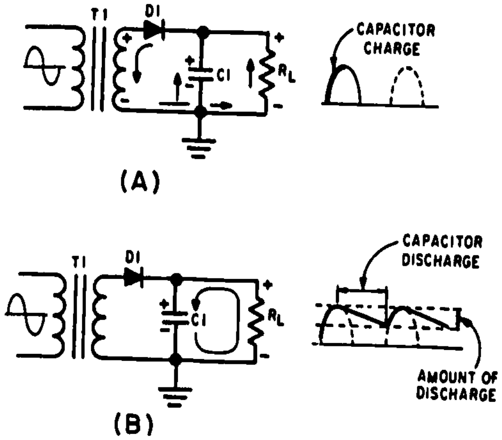
Fig 4-16. Half-wave rectifier with and without filtering.

The value of the capacitor is fairly large (several microfarads), thus it presents a relatively low reactance to the pulsating current and it stores a substantial charge. The rate of charge for the capacitor is limited only by the resistance of the conducting diode which is relatively low. Therefore, the RC charge time of the circuit is relatively short. As a result, when the pulsating voltage is first applied to the circuit, the capacitor charges rapidly and almost reaches the peak value of the rectified voltage within the first few cycles. The capacitor attempts to charge to the peak value of the rectified voltage anytime a diode is conducting, and tends to retain its charge when the rectifier output falls to zero. (The capacitor cannot discharge immediately.) The capacitor slowly discharges through the load resistance (R_L) during the time the rectifier is nonconducting.

The rate of discharge of the capacitor is determined by the value of capacitance and the value of the load resistance. If the capacitance and load-resistance values are large, the RC discharge time for the circuit is relatively long.

A comparison of the wave forms shown in figure 4-16 illustrates that the addition of C1 to the circuit results in an increase in the average of the output voltage (E_{avg}) and a reduction in the amplitude of the ripple component (E_r) which is normally present across the load resistance.

Now, let's consider a complete cycle of operation using a half-wave rectifier, a capacitive filter (C1), and a load resistor (R_L). As shown in view (A) of figure 4-17, the capacitive filter (C1) is assumed to be large enough to ensure a small reactance to the pulsating rectified current. The resistance of R_L is assumed to be much greater than the reactance of C1 at the input frequency. When the circuit is energized, the diode conducts on the positive half cycle and current flows through the circuit, allowing C1 to charge. C1 will charge to approximately the peak value of the input voltage. (The charge is less than the peak value because of the voltage drop across the diode (D1)). In view (A) of the figure, the charge on C1 is indicated by the heavy solid line on the wave form. As illustrated in view (B), the diode cannot conduct on the negative half cycle because the anode of D1 is negative in respect to the cathode. During this interval, C1 discharges through the load resistor (R_L). The discharge of C1 produces the downward slope as indicated by the solid line on the wave form in view (B). In contrast to the abrupt fall of the applied a.c. voltage from peak value to zero, the voltage across C1 (and thus across R_L) during the discharge period gradually decreases until the time of the next half cycle of rectifier operation. Keep in mind that for good filtering, the filter capacitor should charge up as fast as possible and discharge as little as possible.



A. Positive half-cycle
B. Negative half-cycle

Fig 4-17. Capacitor filter circuit (positive and negative half cycles).

Since practical values of C1 and R_L ensure a more or less gradual decrease of the discharge voltage, a substantial charge maintains on the capacitor at the time of the next half cycle of operation. As a result, no current can flow through the diode until the rising a.c. input voltage at the anode of the diode exceeds the voltage of the charge remaining on C1. The charge on C1 is the cathode potential of the diode. When the potential on the anode exceeds the potential on the cathode (the charge on C1), the diode again conducts, and C1 begins to charge to approximately the peak value of the applied voltage.

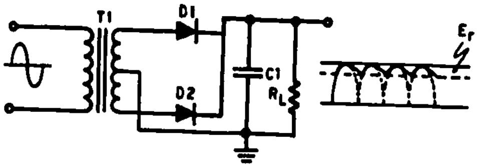


Fig 4-18. Full-wave rectifier (with capacitor filter).

After the capacitor has charged to its peak value, the diode will cut off and the capacitor will start to discharge. Since the fall of the a.c. input voltage on the anode is considerably more rapid than the decrease on the capacitor voltage, the cathode quickly becomes more positive than the anode, and the diode ceases to conduct.

Operation of the simple capacitor filter using a full-wave rectifier is basically the same as that discussed for the half-wave rectifier. Referring to figure 4-18, you should notice that because one of the diodes is always conducting on either alternation, the filter capacitor charges and discharges during each half cycle. (Note that each diode conducts only for that portion of time when the peak secondary voltage is greater than the charge across the capacitor.)

Another thing to keep in mind is that the ripple component (E_r) of the output voltage is an a.c. voltage and the average output voltage (E_{avg}) is the d.c. component of the output. Since the filter capacitor offers a relatively low impedance to a.c., the majority of the a.c. component flows through the filter capacitor. The a.c. component is, therefore, bypassed (shunted) around the load resistance, and the entire d.c. component (or E_{avg}) flows through the load resistance. This statement can be clarified by using the formula for X_C in a half-wave and full-wave rectifier. First, you must establish some values for the circuit.

HALF-WAVE RECTIFIER

FREQUENCY AT RECTIFIER

OUTPUT: 60 Hz

VALUE OF FILTER CAPACITOR:

30μ F

LOAD RESISTANCE (R_L): 10 kΩ

$$X_C = \frac{1}{2\pi fC}$$

$$X_C = \frac{.159}{fC}$$

$$X_C = \frac{.159}{60 \times .000030}$$

$$X_C = \frac{.159}{.0018}$$

$$X_C = 88.3\Omega$$

FULL-WAVE RECTIFIER

FREQUENCY AT RECTIFIER

OUTPUT: 120 Hz

VALUE OF FILTER CAPACITOR:

30μ F

LOAD RESISTANCE (R_L): 10 k

$$X_C = \frac{1}{2\pi fC}$$

$$X_C = \frac{.159}{fC}$$

$$X_C = \frac{.159}{120 \times .000030}$$

$$X_C = \frac{.159}{.036}$$

$$X_C = 44.16\Omega$$

As you can see from the calculations, by doubling the frequency of the rectifier, you reduce the impedance of the capacitor by one-half. This allows the a.c. component to pass through the capacitor more easily. As a result, a full-wave rectifier output is much easier to filter than that of a half-wave rectifier. Remember, the smaller the X_C of the filter capacitor in respect to the load resistance, the better the filtering action since the largest possible capacitor will provide the best filtering. Remember, also, that the load resistance is an important consideration. If load resistance is made small, the load current increases, and the average value of output voltage (E_{avg}) decreases. The RC discharge time constant is a direct function of the value of the load resistance; therefore, the rate of capacitor voltage discharge is a direct function of the current through the load. The greater load current, the more rapid the discharge of the capacitor, and the lower the average value of output voltage. For this reason, the simple capacitive filter is seldom used with rectifier circuits that must supply a relatively large load current. Using the simple capacitive filter in conjunction with a full-wave or bridge rectifier provides improved filtering because the increased ripple frequency decreases the capacitive reactance of the filter capacitor.

$$X_C = \frac{1}{2\pi fC}$$

The LC CHOKE-INPUT FILTER is used primarily in power supplies where voltage regulation is important and where the output current is relatively high and subject to varying load conditions. The filter is used in high power applications such as those found in radars and communication transmitters.

Notice in figure 4-19 that this filter consists of an input inductor (L1), or filter choke, and an output filter capacitor (C1). Inductor L1 is placed at the input to the filter and is in series with the output of the rectifier circuit. Since the action of an inductor is to oppose any change in current flow, the inductor tends to keep a constant current flowing to the load throughout the complete cycle of the applied voltage. As a result, the output voltage never reaches the peak value of the applied voltage. Instead, the output voltage approximates the average value of the rectified input to the filter, as shown in the figure. The reactance of the inductor (X_L) reduces the amplitude of ripple voltage without reducing the d.c. output voltage by an appreciable amount. (The d.c. resistance of the inductor is just a few ohms.)

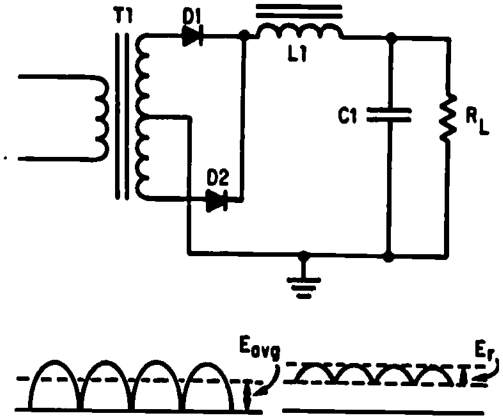
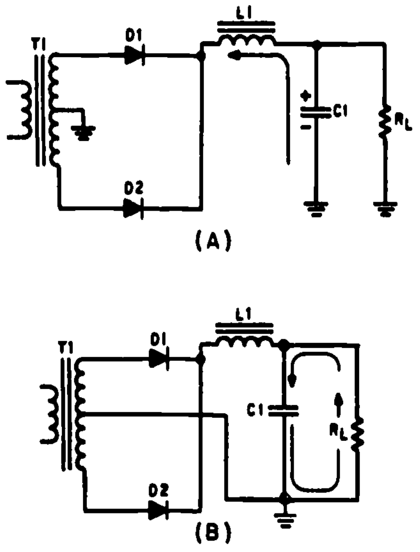


Fig 4-19. LC choke-input filter.

The shunt capacitor (C1) charges and discharges at the ripple frequency rate, but the amplitude of the ripple voltage (E_r) is relatively small because the inductor (L1) tends to keep a constant current flowing from the rectifier circuit to the load. In addition, the reactance of the shunt capacitor (X_C) presents a low impedance to the ripple component existing at the output of the filter, and thus shunts the ripple component around the load. The capacitor attempts to hold the output voltage relatively constant at the average value of the voltage.

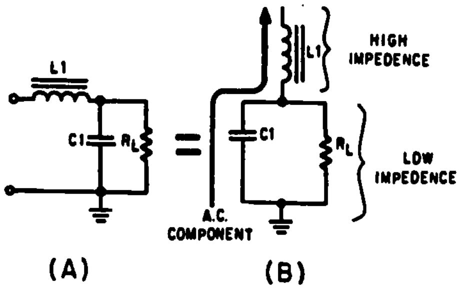
The value of the filter capacitor (C1) must be relatively large to present a low opposition (X_C) to the pulsating current and to store a substantial charge. The rate of the charge for the capacitor is limited by the low impedance of the a.c. source (the transformer), the small resistance of the diode, and the counterelectromotive force (cemf) developed by the coil. Therefore, the RC charge time constant is short compared to its discharge time. (This comparison in RC charge and discharge paths is illustrated in views (A) and (B) of figure 4-20.) Consequently, when the pulsating voltage is first applied to the LC choke-input filter, the inductor (L1) produces a cemf which opposes the constantly increasing input voltage. The net result is to prevent effectively the rapid charging of the filter capacitor (C1). Thus, instead of reaching the peak value of the input voltage, C1 only charges to the average value of the input voltage. After the input voltage reaches its peak and decreases sufficiently, the capacitor (C1) attempts to discharge through the load resistance (R_L). C1 will only partially discharge, as indicated in view (B) of the figure, because of its relatively long discharge time constant. The larger the value of the filter capacitor, the better the filtering action. However, because of physical size, there is a practical limitation to the maximum value of the capacitor.



A. Charge Path
B. Discharge Path

Fig 4-20. LC choke-input filter (charge and discharge paths).

The inductor (also referred to as the filter choke or coil) serves to maintain the current flow to the filter output (R_L) at a nearly constant level during the charge and discharge periods of the filter capacitor. The inductor (L1) and the capacitor (C1) form a voltage divider for the a.c. component (ripple) of the applied input voltage. This is shown in views (A) and (B) of figure 4-21. As far as the ripple component is concerned, the inductor offers a high impedance (Z) and the capacitor offers a low impedance (view B). As a result, the ripple component (E_r) appearing across the load resistance is greatly attenuated (reduced). The inductance of the filter choke opposes change in the value of the current flowing through it; therefore, the average value of the voltage produced across the capacitor contains a much smaller value of ripple component (E_r) than the value of ripple produced across the choke.



A. LC choke-input filter
B. Equivalent circuit

Fig 4-21. LC choke-input filter.

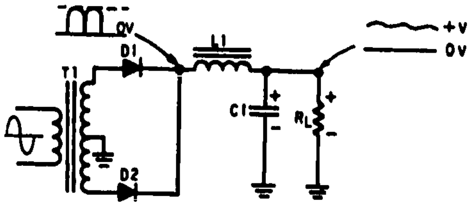


Fig 4-22. Filtering action of the LC choke-input filter.

Now look at figure 4-22 which illustrates a complete cycle of operation for a full-wave rectifier circuit used to supply the input voltage to the filter. The rectifier voltage is developed across the capacitor (C1). The ripple voltage at the output of the filter is the alternating component of the input voltage reduced in amplitude by the filter section. Each time the anode of a diode goes positive with respect to the cathode, the diode conducts the C1 charges. Conduction occurs twice during each cycle for a full-wave rectifier. For a 60-hertz supply, this produces a 120-hertz ripple voltage. Although the diodes alternate (one conducts while the other is nonconducting), the filter input voltage is not steady. As the anode voltage of the conducting diode increases (on the positive half of the cycle), capacitor C1 charges - the charge being limited by the impedance of the secondary transformer winding, the diode's forward (cathode-to-anode) resistance, and the counterelectromotive force developed by the choke. During the nonconducting interval (when the anode voltage drops below the capacitor charge voltage), C1 discharges through the load resistor (R_L). The components in the discharge path have a long time constant; thus, C1 discharges more slowly than it charges.

The choke (L1) is usually a large value, from 1 to 20 henries, and offers a large inductive reactance to the 120-hertz ripple component produced by the rectifier. Therefore, the effect that L1 has on the charging of the capacitor (C1) must be considered. Since L1 is connected in series with the parallel branch consisting of C1 and R_L, a division of the ripple (a.c.) voltage and the output (d.c.) voltage occurs. The greater the impedance of the choke, the less the ripple voltage that appears across C1 and the output. The d.c. output voltage is fixed mainly by the d.c. resistance of the choke.

Now that you have read how the LC choke-input filter functions, it will be discussed with actual component values applied. For simplicity, the input frequency at the primary of the transformer will be 117 volts 60 hertz. Both half-wave and full-wave rectifier circuits will be used to provide the input to the filter.

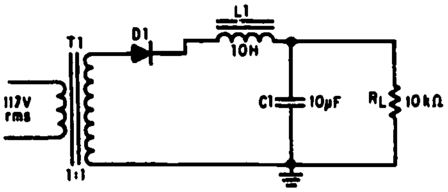


Fig 4-23. Half-wave rectifier with an LC choke-input filter.

Starting with the half-wave configuration shown in figure 4-23, the basic parameters are: With 117 volts a.c. rms applied to the T1 primary, 165 volts a.c. peak-to-peak is available at the secondary [(117 V) x (1.414) = 165 V]. You should recall that the ripple frequency of this half-wave rectifier is 60 hertz. Therefore, the capacitive reactance of C1 is:

$$X_C = \frac{1}{2\pi fC}$$

$$X_C = \frac{1}{(2)\ (3.14)\ (60)\ (10)\ (10^{-6})}$$

$$X_C = \frac{(1)\ (10^6)}{3768}$$

$$X_C = 265 \text{ ohms}$$

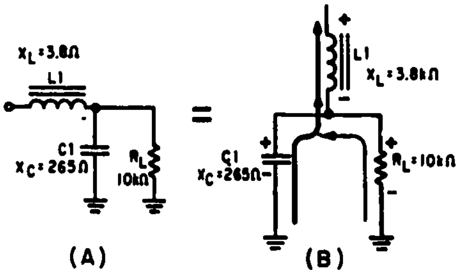
This means that the capacitor (C1) offers 265 ohms of opposition to the ripple current. Note, however, that the capacitor offers an infinite impedance to direct current. The inductive reactance of L1 is:

$$X_L = 2\pi fL$$

$$X_L = (2)\ (3.14)\ (60)\ (10)$$

$$X_L = 3.8 \text{ kilohms}$$

The above calculation shows that L1 offers a relatively high opposition (3.8 kilohms) to the ripple in comparison to the opposition offered by C1 (265 ohms). Thus, more ripple voltage will be dropped across L1 than across C1. In addition, the impedance of C1 (265 ohms) is relatively low in respect to the resistance of the load (10 kilohms). Therefore, more ripple current flows through C1 than the load. In other words, C1 shunts most of the a.c. component around the load.



A. LC choke-input filter
B. Equivalent circuit

Fig 4-24. A.c. component in an LC choke-input filter.

Let's go a step further and redraw the filter circuit so that you can see the voltage divider action. Refer to view (A) of figure 4-24. Remember, the 165 volts peak-to-peak 60 hertz provided by the rectifier consists of both an a.c. and a d.c. component. This first discussion will be about the a.c. component. From the figure, you see that the capacitor (C1) offers the least opposition (265 ohms) to the a.c. component; therefore, the greatest amount of a.c. will flow through C1. (The heavy line indicates the a.c. current flow through the capacitor.) Thus, the capacitor bypasses, or shunts, most of the a.c. around the load.

By combining the X_C of C1 and the resistance of R_L into an equivalent circuit (view B), you will have a total resistance of 258 ohms.

As a formula:

$$RT = \frac{(R1)\ (R2)}{R1 + R2}$$

You now have a voltage divider as illustrated in figure 4-25. You should see that because of the impedance ratios, a large amount of ripple voltage is dropped across L1, and a substantially smaller amount is dropped across C1 and R_L. You can further increase the ripple voltage across L1 by increasing the inductance ($X_L = 2\pi fL$).

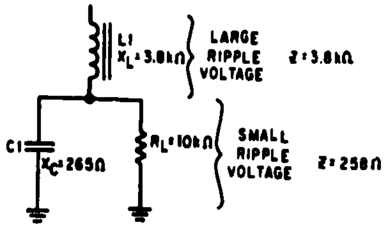


Fig 4-25. Equivalent circuit of an LC choke-input filter.

Now let's discuss the d.c. component of the applied voltage. Remember, a capacitor offers an infinite (∞) impedance to the flow of direct current. The d.c. component, therefore, must flow through R_L and L1. As far as the d.c. is concerned, the capacitor does not exist. The coil and the load are, therefore, in series with each other. The d.c. resistance of a filter choke is very low (50 ohms average). Consequently, most of the d.c. component is developed across the load and a very small amount of the d.c. voltage is dropped across the coil, as shown in figure 4-26.

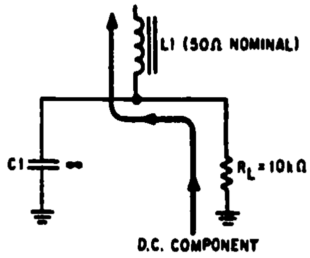


Fig 4-26. D.C. component in an LC choke-input filter.

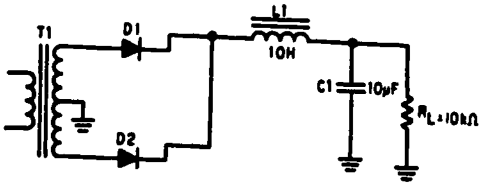


Fig 4-27. Full-wave rectifier with an LC choke-input filter.

As you may have noticed, both the a.c. and the d.c. components flow through L1. Because it is frequency sensitive, the coil provides a large resistance to a.c. and a small resistance to d.c. In other words, the coil opposes any change in current. This property makes the coil a highly desirable filter component. Note that the filtering action of the LC choke-input filter is improved when the filter is used in conjunction with a full-wave rectifier, as shown in figure 4-27. This is due to the decrease in the X_C of the filter capacitor and the increase in the X_L of the choke. Remember, ripple frequency of a full-wave rectifier is twice that of a half-wave rectifier. For a 60-hertz input, the ripple will be 120 hertz. The X_C of C1 and the X_L of L1 is calculated as follows:

$$X_C = \frac{1}{2 \pi fC}$$

$$X_C = \frac{1}{(2)\ (3.14)\ (60)\ (10)\ (10^{-6})}$$

$$X_C = \frac{(1)\ (10^{6})}{7536}$$

$$X_C = 132.5 \text{ ohms}$$

$$X_L = 2 \pi fL$$

$$X_L = (2)\ (3.14)\ (120)\ (10)$$

$$X_L = 7.5 \text{ Kilohms}$$

When the X_C of a filter capacitor is decreased, it provides less opposition to the flow of a.c. The greater the a.c. flow through the capacitor, the lower the flow through the load. Conversely, the larger the X_L of the choke, the greater the amount of a.c. ripple developed across the choke; consequently, less ripple is developed across the load and better filtering is obtained.

Failure analysis of an LC choke-input filter. The filter capacitors are subject to open circuits, short circuits, and excessive leakage; the series inductor is subject to open windings and, occasionally, shorted turns or a short circuit to the core.

The filter capacitor in the LC choke-input filter circuit is not subject to extreme voltage surges because of the protection offered by the inductor; however, the capacitor can become open, leaky, or shorted.

Shorter turns in the choke may reduce the value of inductance below the critical value. This will result in excessive peak-rectifier current, accompanied by an abnormally high output voltage, excessive ripple amplitude, and poor voltage regulation.

A choke winding that is open, or a choke winding which is shorted to the core will result in a no-output condition. A choke winding which is shorted to the core may cause overheating of the rectifier element(s) and blown fuses.

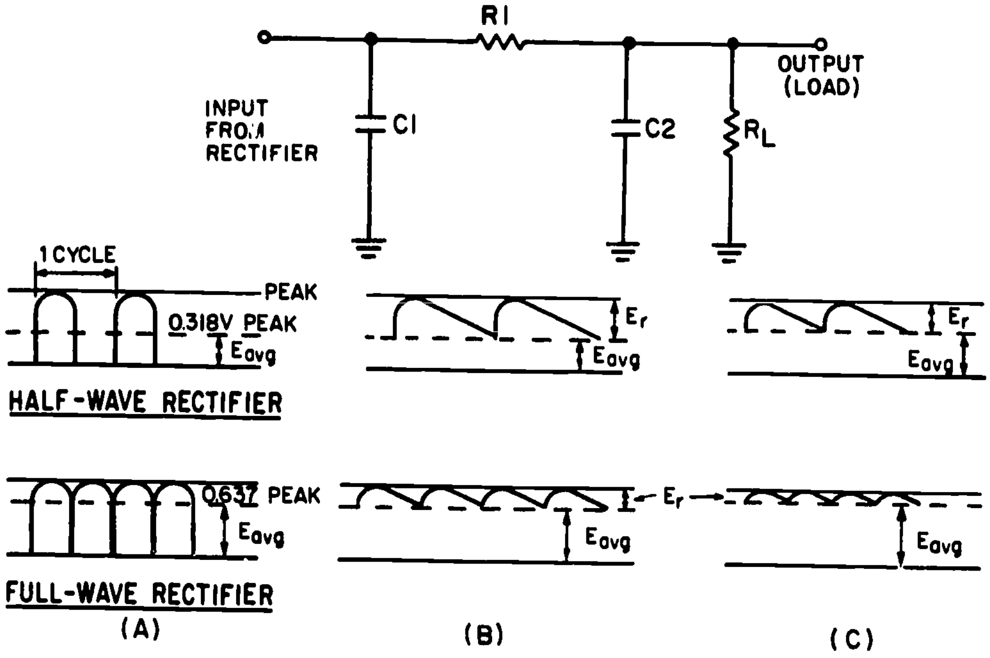
With the supply voltage removed from the input to the filter circuit, one terminal of the capacitor can be disconnected from the circuit. The capacitor should be checked with a capacitance analyzer to determine its capacitance and leakage resistance. When the capacitor is electrolytic, you must use the correct polarity at all times. A decrease in capacitance or losses within the capacitor can decrease the efficiency of the filter and can produce excessive ripple amplitude.

Resistor-capacitor (RC) filters. The RC capacitor-input filter is limited to applications in which the load current is small. This type of filter is used in power supplies where the load current is constant and voltage regulation is not necessary. For example, RC filters are used in high-voltage power supplies for cathode-ray tubes and decoupling networks for multistage amplifiers.

Figure 4-28 shows an RC capacitor-input filter and associated wave forms. Both half-wave and full-wave rectifiers are used to provide the inputs. The wave forms shown in view (A) of the figure represent the unfiltered output from a typical rectifier circuit. Note that the dashed lines in view (A) indicate the average value of output voltage (E_{avg}) for the half-wave rectifier. The average output voltage (E_{avg}) is less than half (approximately 0.318) the amplitude of the voltage peaks. The average value of output voltage (E_{avg}) for the full-wave rectifier is greater than half (approximately 0.637), but is still much less than the peak amplitude of the rectifier-output wave form. With no filter circuit connected across the output of the rectifier circuit (unfiltered), the wave form has a large value of pulsating component (ripple) as compared to the average (or d.c.) component.

The RC filter in figure 4-28 consists of an input filter capacitor (C1), a series resistor (R1), and an output filter capacitor (C2). (This filter is sometimes referred to as an RC pi-section filter because its schematic symbol resembles the Greek letter π).

The single capacitor filter is suitable for many noncritical, low-current applications. However, when the load resistance is very low or when the percent of ripple must be held to an absolute minimum, the capacitor value required must be extremely large. While electrolytic capacitors are available in sizes up to 10,000 microfarads or greater, the large sizes are quite expensive. A more practical approach is to use a more sophisticated filter that can do the same job but that has lower capacitor values, such as the RC filter.



A. Unfiltered output voltage from rectifier
B. Voltage across capacitor C1
C. Voltage across capacitor C2

Fig 4-28. RC filter and waveforms.

Views (A), (B), and (C) of figure 4-28 show the output wave forms of a half-wave and full-wave rectifier. Each wave form is shown with an RC filter connected across the output. The following explanation of how a filter works will show you than an RC filter of this type does a much better job than the single capacitor filter.

C1 performs exactly the same function as it did in the single capacitor filter. It is used to reduce the percentage of ripple to a relatively low value. Thus, the voltage across C1 might consist of an average d.c. value of +100 volts with a ripple voltage of 10 volts peak-to-peak. This voltage is passed on to the R1-C2 network, which reduces the ripple even further.

C2 offers an infinite impedance (resistance) to the d.c. component of the output voltage. Thus, the d.c. voltage is passed to the load, but reduced in value by the amount of the voltage drop across R1. However, R1 is generally small compared to the load resistance. Therefore, the drop in the d.c. voltage by R1 is not a drawback.

Component values are designed so that the resistance of R1 is much greater than the reactance (X_C) of C2 at the ripple frequency. C2 offers a very low impedance to the a.c. ripple frequency. Thus, the a.c. ripple senses a voltage divider consisting of R1 and C2 between the output of the rectifier and ground. Therefore, most of the ripple voltage is dropped across R1. Only a trace of the ripple voltage can be seen across C2 and the load. In extreme cases where the ripple must be held to an absolute minimum, a second stage of RC filtering can be added. In practice, the second stage is rarely required. The RC filter is extremely popular because smaller capacitors can be used with good results.

The RC filter has some disadvantages. First, the voltage drop across R1 takes voltage away from the load. Second, power is wasted in R1 and is dissipated in the form of unwanted heat. Finally, if the load resistance changes, the voltage across the load will change. Even so, the advantages of the RC filter overshadow these disadvantages in many cases.

Failure analysis of the resistor-capacitor (RC) filter. The shunt capacitors (C1 and C2) are subject to an open circuit, a short circuit, or excessive leakage. The series filter resistor (R1) is subject to changes in value and occasionally opens. Any of these troubles can be easily detected.

The input capacitor (C1) has the greatest pulsating voltage applied to it and is the most susceptible to voltage surges. As a result, the input capacitor is frequently subject to voltage breakdown and shorting. The remaining shunt capacitor (C2) in the filter circuit is not subject to voltage surges because of the protection offered by the series filter resistor (R1). However, a shunt capacitor can become open, leaky, or shorted.

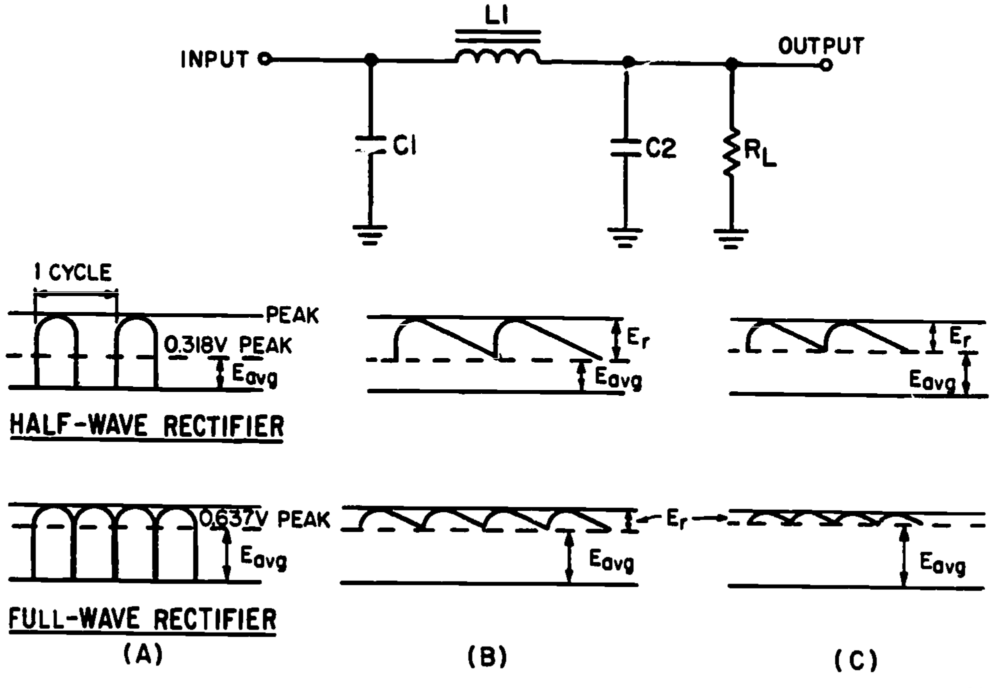
A shorted capacitor or an open filter resistor results in a no-output indication. An open filter resistor results in an abnormally high d.c. voltage at the input to the filter and no voltage at the output of the filter. Leaky capacitors or filter resistors that have lost their effectiveness, or filter resistors that have decreased in value, result in an excessive ripple amplitude in the output of the supply.

The LC CAPACITOR-INPUT FILTER is one of the most commonly used filters. This type of filter is used primarily in radio receivers, small audio amplifier power supplies, and in any type of power supply where the output current is low and the load current is relatively constant.

Figure 4-29 shows an LC capacitor-input filter and associated wave forms. Both half-wave and full-wave rectifier circuits are used to provide the input. The wave forms shown in view (A) of the figure represent the unfiltered output from a typical rectifier circuit. Note that the average value of output voltage (E_{avg}), indicated by the dashed lines, for the half-wave rectifier is less than half the amplitude of the voltage peaks. The average value of output voltage (E_{avg}) for the full-wave rectifier is greater than half, but is still much less than the peak amplitude of the rectifier-output waveform. With no filter connected across the output of the rectifier circuit (which results in unfiltered output voltage), the wave form has a large value of pulsating component (ripple) as compared to the average (or d.c.) component.

C1 reduces the ripple to a relatively low level (view B). L1 and C2 form the LC filter which reduces the ripple even further. L1 is a large value iron-core inductor (choke). L1 has a high value of inductance and, therefore, a high value of X_L which offers a high reactance to the ripple frequency. At the same time, C2 offers a very low reactance to a.c. ripple. L1 and C2 form an a.c. voltage divider and, because the reactance of L1 is much higher than that of C2, most of the ripple voltage is dropped across L1. Only a slight trace of ripple appears across C2 and the load (view C).

While the L1-C2 network greatly reduces a.c. ripple, it has little effect on d.c. ripple. You should recall that an inductor offers no reactance to d.c. The only opposition to current flow is the resistance of the wire in the choke. Generally, this resistance is very low and the d.c. voltage drop across the coil is minimal. Thus, the LC filter overcomes the disadvantages of the RC filter.



A. Unfiltered output voltage from rectifier B. Voltage across capacitor C1 C. Voltage across capacitor C2

Fig 4-29. LC filter and waveforms.

Aside from the voltage divider effect, the inductor improves filtering in another way. You should recall that an inductor resists changes in the magnitude of the current flowing through it. Consequently, when the inductor is placed in series with the load, the inductor maintains steady current. In turn, this helps the voltage across the load remain constant when the size of the components is a factor.

The LC filter provides good filtering action over a wide range of currents. The capacitor filters best when the load is drawing little current. Thus, the capacitor discharges very slowly and the output voltage remains almost constant. On the other hand, the inductor filters best when the current is highest. The complementary nature of these two components ensures that good filtering will occur over a wide range of currents.

The LC filter has two disadvantages. First, it is more expensive than the RC filter because an iron-core choke costs more than a resistor. The second disadvantage is size. The iron-core choke is bulky and heavy, a fact which may render the LC filter unsuitable for many applications.

Failure analysis of the LC capacitor-input filter. Shunt capacitors are subject to open circuits, short circuits, and excessive leakage; series inductors are subject to open windings and occasionally shorted turns or a short circuit to the core.

The input capacitor (C1), which has the greatest pulsating voltage applied to it, is the most susceptible to voltage surges, and has a generally higher average voltage applied. As a result, the input capacitor is frequently subject to voltage breakdown and shorting. The output capacitor (C2) is not as susceptible to voltage surges because of the series protection offered by the series inductor (L1), but the capacitor can become open, leaky, or shorted.

A shorted capacitor, an open filter choke, or a choke winding which is shorted to the core, results in a no-output indication. A shorted capacitor, depending on the magnitude of the short, may cause a shorted rectifier, transformer, or filter choke, and may result in a blown fuse in the primary of the transformer. An open filter choke results in an abnormally high d.c. voltage at the input to the filter and no voltage at the output of the filter. A leaky or open capacitor in the filter circuit results in a low d.c. output voltage. This condition is generally accompanied by an excessive ripple amplitude. Shorted turns in the winding of a filter choke reduce the effective inductance of the choke and decrease its filtering efficiency. As a result, the ripple amplitude increases.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What are the four basic sections of a power supply?

   a. ______________________________

   b. ______________________________

   c. ______________________________

   d. ______________________________

2. What is the purpose of the rectifier section?

   ____

   ____

3. What is the purpose of the filter section?

   ____

   ____

4. What is the purpose of the regulator section?

   ____

   ____

5. What is the name of the simplest type of rectifier which uses one diode?

6. If the output of a half-wave rectifier is 50 volts peak, what is the average voltage?

7. In addition to stepping up or stepping down the input line voltage, what additional purpose does the transformer serve?

8. What was the major factor that led to the development of the full-wave rectifier?

9. What is the ripple frequency of a full-wave rectifier with an input frequency of 60 hertz?

10. What is the average voltage (E_{avg}) output of a full-wave rectifier with an output of 100 volts peak?

11. What is the main disadvantage of a conventional full-wave rectifier?

12. What main advantage does a bridge rectifier have over a conventional full-wave rectifier?

13. If you increase the value of the capacitor will the X_C increase or decrease?

14. What is the most basic type of filter?

15. In a capacitor filter, is the capacitor in series or in parallel with the load?

16. Is filtering better at a high frequency or at a low frequency?

17. Does a filter circuit increase or decrease the overage output voltage?

18. What determines the rate of discharge of the capacitor in a filter circuit?

19. Does low ripple voltage indicate good or bad filtering?

20. Is a full-wave rectifier output easier to filter than that of a half-wave rectifier?

21. In a LC choke-input filter, what prevents the rapid charging of the capacitor?

22. What is the range of values usually chosen for a choke?

23. If the impedance of the choke is increased, will the ripple frequency increase or decrease?

24. Why is the use of large value capacitors in filter circuits discouraged?

25. When is a second RC filter stage used?

26. What is the most commonly used filter today?

27. What are the two main disadvantages of an LC capacitor filter?

a. ____________________

b. ____________________

Work Unit 4-2. VOLTAGE REGULATION AND MULTIPLIERS

DESCRIBE THE OPERATION OF THE VARIOUS VOLTAGE AND CURRENT REGULATORS IN A POWER SUPPLY.

DESCRIBE THE OPERATION OF THE VARIOUS TYPES OF VOLTAGE MULTIPLIERS.

TRACE THE FLOW OF A.C. AND D.C. IN A POWER SUPPLY, FROM THE A.C. INPUT TO THE D.C. OUTPUT ON A SCHEMATIC DIAGRAM.

Ideally, the output of most power supplies should be a constant voltage. Unfortunately, this is difficult to achieve. There are two factors which can cause the output voltage to change. First, the a.c. line voltage is not constant. The so-called 115 volts a.c. can vary from about 105 volts a.c. to 125 volts a.c. This means that the peak a.c. voltage to which the rectifier responds can vary from about 148 volts to 177 volts. The a.c. line voltage alone can be responsible for nearly a 20 percent change in the d.c. output voltage. The second factor that can change the d.c. output voltage is a change in the load resistance. In complex electronic equipment, the load can change as circuits are switched in and out. In a television receiver, the load on a particular power supply may depend on the brightness of the screen, the control settings, or even the channel selected.

These variations in load resistance tend to change the applied d.c. voltage because the power supply has a fixed internal resistance. If the load resistance decreases, the internal resistance of the power supply drops more voltage. This causes a decrease in the voltage across the load.

Many circuits are designed to operate with a particular supply voltage. When the supply voltage changes, the operation of the circuit may be adversely affected. Consequently, some types of equipment must have power supplies which produce the same output voltage regardless of changes in the load resistance or changes in the a.c. line voltage. This constant output voltage may be achieved by adding a circuit called the VOLTAGE REGULATOR at the output of the filter. There are many different types of regulators in use today and to discuss all of them would be beyond the scope of this course.

LOAD REGULATION

A commonly used FIGURE OF MERIT for a power supply is its PERCENT OF REGULATION. The figure of merit gives us an indication of how much the output voltage changes over a range of load resistance values. The percent of regulation aids in the determination of the type of load regulation needed. Percent of regulation is determined by the equation:

$$\text{Percent of Regulation} = \frac{E_{\text{no load}} - E_{\text{full load}}}{E_{\text{full load}}} \times 100$$

This equation compares the change in output voltage at the two loading extremes to the voltage produced at full loading. For example, assume that a power supply produces 12 volts when the load current is zero. If the output voltage drops to 10 volts when full load current flows, the percent of regulation is:

$$\text{Percent of Regulation} = \frac{E_{\text{no load}} - E_{\text{full load}}}{E_{\text{full load}}} \times 100$$

$$= \frac{12 - 10}{10} \times 100$$

$$= \frac{2}{10} \times 100$$

$$= 20\%$$

Ideally, the output voltage should not change over the full range of operation. That is, a 12-volt power supply should produce 12 volts at no load, at full load, and at all points in between. In this case, the percent of regulation would be:

$$\text{Percent of Regulation} = \frac{E_{\text{no load}} - E_{\text{full load}}}{E_{\text{full load}}} \times 100$$

$$\text{Percent of Regulation} = \frac{12 - 12}{12} \times 100$$

$$\text{Percent of Regulation} = \frac{0}{12} \times 100$$

$$\text{Percent of Regulation} = 0\%$$

Thus, zero-percent load regulation is the ideal situation. It means that the output voltage is constant under all load conditions. While you should strive for zero-percent load regulation, in practical circuits you must settle for something less ideal. Even so, by using a voltage regulator, you can hold the percent of regulation to a very low value.

REGULATORS

You should know that the output of a power supply varies with changes in input voltage and circuit load current requirements. Because many electronic equipment require operating voltages and currents which must remain constant, some form of regulation is necessary. Circuits which maintain power supply voltages or current outputs within specified limits, or tolerances, are called REGULATORS. They are designated as d.c. voltage or d.c. current regulators, depending on their specific application.

Voltage regulator circuits are additions to basic power supply circuits which are made up of rectifier and filter sections (figure 4-30). The purpose of the voltage regulator is to provide an output voltage with little or no variation. Regulator circuits sense changes in output voltages and compensate for the changes. Regulators that maintain voltages within plus or minus (±) 0.1 percent are quite common.

Series or shunt voltage regulators. There are two basic types of voltage regulators which are classified as either SERIES or SHUNT, depending on the location or position of the regulating element(s) in relation to the circuit load resistance. Figure 4-31 illustrates these two basic types of voltage regulators. In actual practice the circuitry of regulating devices may be quite complex. Broken lines have been used in the figure to highlight the differences between the series and shunt regulators.

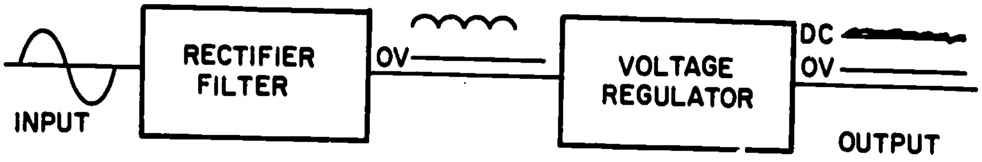


Fig 4-30. Block diagram of a power supply and regulator.

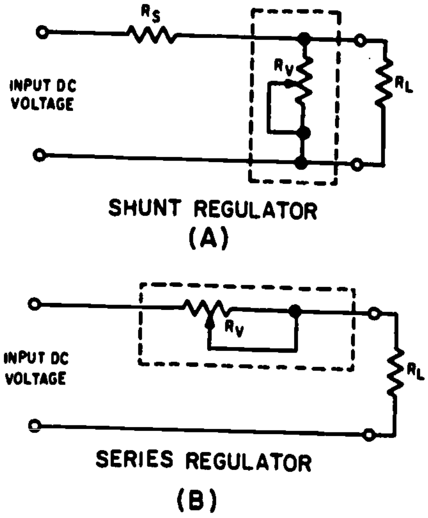


Fig 4-31. Simple series and shunt regulators.

The schematic drawing in view (A) is that of a shunt-type regulator. It is called a shunt-type regulator because the regulating device is connected in parallel with the load resistance. The schematic drawing in view (B) is that of a series regulator. It is called a series regulator because the regulating device is connected in series with the load resistance. Figure 4-32 illustrates the principle of series voltage regulation. As you study the figure, notice that the regulator is in series with the load resistance (R_L) and that the fixed resistor (R_S) is in series with the load resistance.

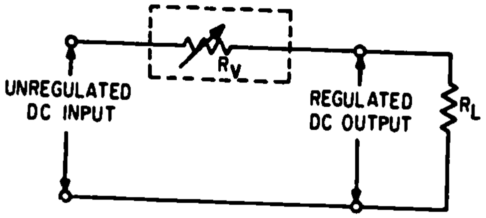


Fig 4-32. Series voltage regulator.

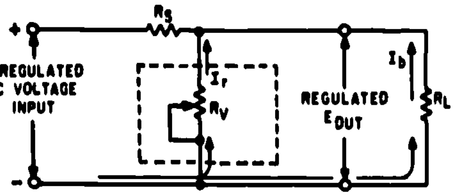


Fig 4-33. Shunt regulator.

You already know the voltage drop across a fixed resistor remains constant unless the current flowing through it varies (increases or decreases). In a shunt regulator, as shown in figure 4-33, output voltage regulation is determined by the current through the parallel resistance of the regulating device (R_V), the load resistance (R_L), and the series resistor (R_S). For now, assume that the circuit is operating under normal conditions, the input is 120 volts d.c., and the desired regulated output is 100 volts d.c. For a 100-volt output to be maintained, 20 volts must be dropped across the series resistor (R_S). If you assume that the value of R_S is 2 ohms, then you must have 10 amperes of current across R_V and R_L. (Remember: $E = IR$.) If the values of the resistance of R_V and R_L are equal, then 5 amperes of current will flow through each resistance (R_V and R_L).

Now, if the load resistance (R_L) increases, the current through R_L will decrease. For example, assume that the current through R_L is now 4 amperes and that the total current across R_S is 9 amperes. With this drop in current, the voltage drop across R_S is 18 volts; consequently, the output of the regulator has increased to 102 volts. At this time, the regulating device (R_V) decreases in resistance, and 6 amperes of current flows through this resistance (R_V). Thus, the total current R_S is once again 10 amperes (6 amperes across R_V; 4 amperes across R_L). Therefore, 20 volts is dropped across R_S causing the output to decrease back to 100 volts. You should know by now that if the load resistance (R_L) increases, the regulating device (R_V) decreases its resistance to compensate for the change. If R_L decreases, the opposite effect occurs and R_V increases.

Now consider the circuit when a decrease in load resistance takes place. When R_L decreases, the current through R_L subsequently increases to 6 amperes. This action causes a total of 11 amperes to flow through R_S which then drops 22 volts. As a result, the output is 98 volts. However, the regulating device (R_V) senses this change and increases its resistance so that less current (4 amperes) flows through R_V. The total current again becomes 10 amperes, and the output is again 100 volts.

From these examples, you should now understand that the shunt regulator maintains the desired output voltage first by sensing the current change in the parallel resistance of the circuit and then by compensating for the change.

Again refer to the schematic shown in figure 4-33 and consider how the voltage regulator operates to compensate for changes in input voltages. You know of course, that the input voltage may vary and that any variation must be compensated for by the regulating device. If an increase in input voltage occurs, the resistance of R_V automatically decreases to maintain the correct voltage division between R_V and R_S. You should see, therefore, that the regulator operates in the opposite way to compensate for a decrease in input voltage.

So far only voltage regulators that use variable resistors have been explained; however, this type of regulation has limitations. Obviously the variable resistor cannot be adjusted rapidly enough to compensate for frequent fluctuations in voltage. Since input voltages fluctuate frequently and rapidly, the variable resistor is not a practical method for voltage regulation. A voltage regulator that operates continuously and automatically to regulate the output voltage without external manipulation is required for practical regulation.

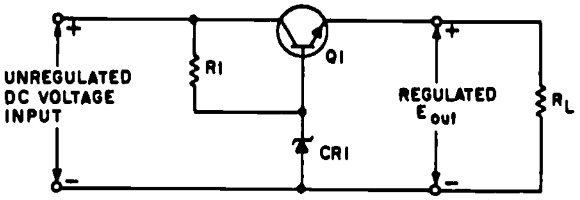


Fig 4-34. Series voltage regulator.

The schematic for a typical series voltage regulator is shown in figure 4-34. Notice that this regulator has a transistor (Q1) in the place of the variable resistor found in figure 4-32. Because the total load current passes through this transistor, it is sometimes called a "pass transistor." Other components which make up the circuit are the current limiting resistor (R1) and the Zener diode (CR1).

Recall that a Zener diode is a diode which blocks current until a specified voltage is applied. Remember also that the applied voltage is called the breakdown, or Zener voltage. Zener diodes are available with different Zener voltages. When the Zener voltage is reached, the Zener diode conducts from its anode to its cathode (with the direction of the arrow).

In this voltage regulator, Q1 has a constant voltage applied to its base. This voltage is often called the reference voltage. As changes in the circuit output voltage occur, they are sensed at the emitter of Q1, producing a corresponding change in the forward bias of the transistor. In other words, Q1 compensates by increasing or decreasing its resistance in order to change the circuit voltage division.

Now, study figure 4-35. Voltages are shown to help you understand how the regulator operates. The Zener used in this regulator is a 15-volt Zener. In this instance, the Zener or breakdown voltage is 15 volts. The Zener establishes the value of the base voltage for Q1. The output voltage will equal the Zener voltage minus a 0.7-volt drop across the forward biased base-emitter junction of Q1, or 14.3 volts. Because the output voltage is 14.3 volts, the voltage drop across Q1 must be 5.7 volts.

Study figure 4-36, view (A), in order to understand what happens when the input voltage exceeds 20 volts. Notice the input and output voltages of 20.1 and 14.4 volts, respectively. The 14.4 output voltage is a momentary deviation, or variation, from the required regulated output voltage of 14.3 and is the result of a rise in the input voltage to 20.1 volts. Since the base voltage of Q1 is held at 15 volts by CR1, the forward bias of Q1 changes to 0.6 volt. Because this bias voltage is less than the normal 0.7 volt, the resistance of Q1 increases, thereby increasing the voltage drop across the transistor to 5.8 volts. The voltage drop restores the output voltage to 14.3 volts. The entire cycle takes only a fraction of a second and, therefore, the change is not visible on an oscilloscope or readily measurable with other standard test equipment.

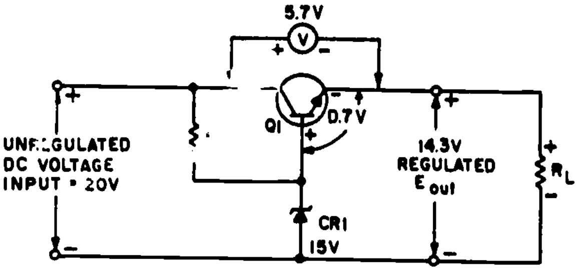
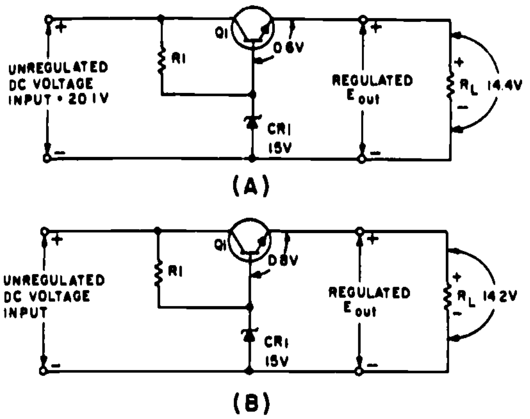


Fig 4-35. Series voltage regulator (with voltages).



A. Increase in output
B. Decrease in output

Fig 4-36. Series voltage regulator.

View (B) is a schematic diagram for the same series voltage regulator with one significant difference. The output voltage is shown as 14.2 volts instead of the desired 14.3 volts. In this case, the load has increased causing a greater voltage drop across R_L to 14.2 volts. When the output decreases, the forward bias of Q1 increases to 0.8 volt because Zener diode CR1 maintains the base voltage of Q1 at 15 volts. This 0.8 volt is the difference between the Zener reference voltage of 15 volts and the momentary output voltage (15 V - 14.2 V = 0.8 V). At this point, the larger forward bias on Q1 causes the resistance of Q1 to decrease, thereby causing the voltage drop across Q1 to return to 5.7 volts. This then causes the output voltage to return to 14.3 volts.

The schematic shown in figure 4-37 is that of a shunt voltage regulator. Notice that Q1 is in parallel with the load. Components of this circuit are identical with those of the series voltage regulator except for the addition of fixed resistor R_S. As you study the schematic, you will see that this resistor is connected in series with the output load resistance. The current limiting resistor (R1) and Zener diode (CR1) provide a constant reference voltage for the base-collector junction of Q1. Notice that the bias of Q1 is determined by the voltage drop across R_S and R1. As you should know, the amount of forward bias across a transistor affects its total resistance. In this case, the voltage drop across R_S is the key to the total circuit operation.

Figure 4-38 is the schematic for a typical shunt-type regulator. Notice that the schematic is identical to the schematic shown in figure 4-37 except that voltages are shown to help you understand the functions of the various components. In the circuit shown, the voltage drop across the Zener diode (CR1) remains constant at 5.6 volts. This means that with a 20-volt input voltage, the voltage drop across R1 is 14.4 volts. With a base-emitter voltage of 0.7 volt, the output voltage is equal to the sum of the voltages across CR1 and the voltage at the base-emitter junction of Q1. In this example, with an output voltage of 6.3 volts and a 20-volt input voltage, the voltage drop across R_S equals 13.7 volts. Study the schematic to understand fully how these voltages are developed. Pay close attention to the voltages shown.

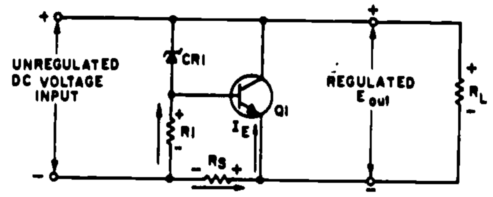


Fig 4-37. Shunt voltage regulator.

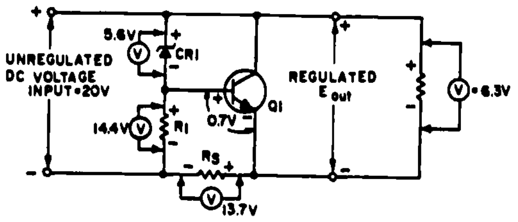
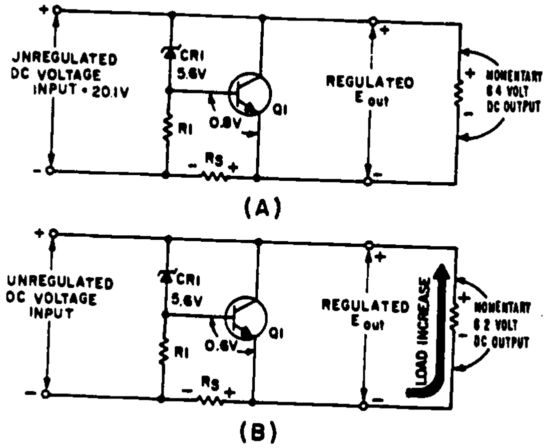


Fig 4-38. Shunt voltage regulator (with voltages).



A. Increase in output voltage
B. Decrease in output voltage

Fig 4-39. Shunt voltage regulator.

Now refer to view (A) of figure 4-39. This figure shows the schematic diagram of the same shunt voltage regulator as that shown in figure 4-38 with an increased input voltage of 20.1 volts. This increases the forward bias on Q1 to 0.8 volt. Recall that the voltage drop across CR1 remains constant at 5.6 volts. Since the output voltage is comprised of the Zener voltage and the base-emitter voltage, the output voltage momentarily increases to 6.4 volts. At this time, the increase in the forward bias of Q1 lowers the resistance of the transistor allowing more current to flow through it. Since this current must also pass through R_S, there is also an increase in the voltage drop across this resistor. The voltage drop across R_S is now 13.8 volts and, therefore, the output voltage is reduced to 6.3 volts. Remember, this change takes place in a fraction of a second.

Study the schematic shown in view (B). Although this schematic is identical to the other shunt voltage schematics previously illustrated and discussed, the output voltage is different. The load current has increased causing a momentary drop in voltage output to 6.2 volts. Recall that the circuit was designed to ensure a constant output voltage of 6.3 volts. Since the output voltage is less than that required, changes occur in the regulator to restore the output to 6.3 volts. Because of the 0.1-volt drop in the output voltage, the forward bias of Q1 is now 0.6 volt. The decrease in the forward bias increases the resistance of the transistor, thereby reducing the current flow through Q1 by the same amount that the load current increased. The current flow through R_s returns to its normal value and restores the output voltage to 6.3 volts.

Current regulators. You should now know how voltage regulators work to provide constant output voltage. In some circuits it may be necessary to regulate the current output. The circuitry which provides a constant current output is called a constant current regulator or just CURRENT REGULATOR. The schematic shown in figure 4-40 is a simplified schematic for a current regulator. The variable resistor shown on the schematic is used to illustrate the concept of current regulation. You should know from your study of voltage regulators that a variable resistor does not respond quickly enough to compensate for the changes. Notice that an ammeter has been included in this circuit to indicate that the circuit shown is that of a current regulator. When the circuit functions properly, the current reading of the ammeter remains constant. In this case the variable resistor (R_V) compensates for changes in the load or d.c. input voltage. Adequate current regulation results in the loss of voltage regulation. Studying the schematic shown, you should recall that any increase in load resistance causes a drop in current. To maintain a constant current flow, the resistance of R_V must be reduced whenever the load resistance increases. This causes the total resistance to remain constant. An increase in the input voltage must be compensated for by an increase in the resistance of R_V, thereby maintaining a constant current flow. The operation of a current regulator is similar to that of a voltage regulator. The basic difference is that one regulates current and the other regulates voltage.

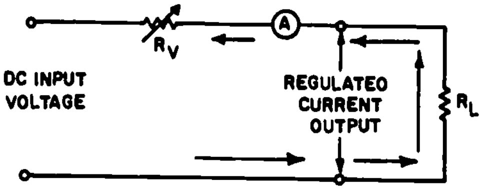


Fig 4-40. Current regulator.

Since use of a variable resistor is not a practical way to control current fluctuation, or variation, a transistor and a Zener diode, together with necessary resistors are used. Recall that the Zener diode provides a constant reference voltage. The schematic shown in figure 4-41 is that of a current regulator circuit. Except for the addition of R1, the circuit shown in the figure is similar to that of a series voltage regulator. The resistor is connected in series with the load and senses any current changes in the load. Notice the voltage drop across R1 and the negative voltage polarity applied to the emitter of Q1. The voltage polarity is a result of current flowing through R1, and this negative voltage opposes the forward bias for Q1; however, since the regulated voltage across CR1 has an opposite polarity, the actual bias of the transistor is the difference between the two voltages. You should see therefore that the purpose of R2 is to function as a current-limiting resistor for the Zener diode.

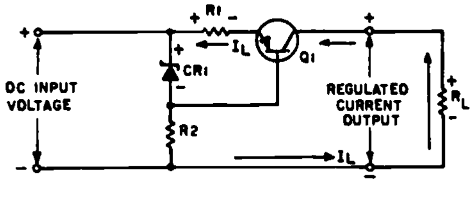


Fig 4-41. Current regulator.

The purpose of a current regulator is to provide a constant current regardless of changes in the input voltage or load current. The schematic shown in figure 4-42 is that of circuit designed to provide a constant current of 400 milliamperes. Voltmeters are shown in the schematic to emphasize the voltage drops across specific components. These voltages will help you understand how the current regulator operates. The voltage drop across the base-emmiter junction of Q1 is 0.6 volt. This voltage is the difference between the Zener voltage and the voltage drop across R1. The 0.6-volt forward bias of Q1 permits proper operation of the transistor. The output voltage across R_L is 6 volts as shown by the voltmeter. With a regulated current output of 400 milliamperes, the transistor resistance (R_{Q1}) is 9 ohms. This can be proved by using Ohm's law and the values shown on the schematic. In this case, current (I) is equal to the voltage drop (E) divided by the resistance (R). Therefore, 12 volts divided by 30 ohms equals 0.4 ampere, or 400 milliamperes.

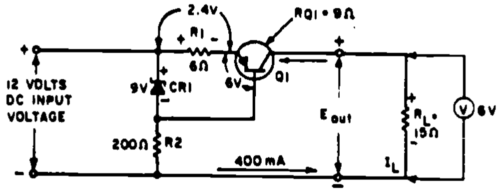


Fig 4-42. Current regulator (with circuit values).

Since you are familiar with the basic current regulating circuitry, let's examine in detail how the various components work to maintain the constant 400-milliamperes output. Refer to the schematic shown in figure 4-43. Remember, a decrease in load resistance causes a corresponding increase in current flow. In the example shown, the load resistance R_L has dropped from 15 ohms to 10 ohms. This results in a larger voltage drop across R1 because of the increased current flow. The voltage drop has increased from 2.4 volts to 2.5 volts. Of course, the voltage drop across CR1 remains constant at 9 volts due to its regulating ability. Because of the increased voltage drop across R1, the forward bias on Q1 is now 0.5 volt. Since the forward bias of Q1 has decreased, the resistance of the transistor increases from 9 ohms to 14 ohms. Notice that the 5-ohm increase in resistance across the transistor corresponds to the 5-ohm decrease in the load resistance. Thus, the total resistance around the outside loop of the circuit remains constant. Since the circuit is a current regulator, you know that output voltage will vary as the regulator maintains a constant current output. In the figure, the voltage output is reduced to 4 volts, which is computed by multiplying current (I) times resistance (R) (400 mA x 10 ohms = 4 volts).

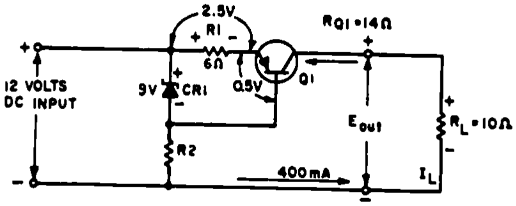


Fig 4-43. Current regulator (with a decrease in R_L).

VOLTAGE MULTIPLIERS

You may already know how a transformer functions to increase or decrease voltage. You may also have learned that a transformer secondary may provide one or several a.c. voltage outputs which may be greater or less than the input voltage. When voltages are stepped up, current is decreased; when voltages are stepped down, current is increased.

Another method for increasing voltages is known as voltage multiplication. VOLTAGE MULTIPLIERS are used primarily to develop high voltages where low current is required. The most common application of the high voltage outputs of voltage multipliers is the anode of cathode-ray tubes (CRT) which are used for radar scope presentations, oscilloscope presentations, or TV picture tubes. The d.c. output of the voltage multiplier ranges from 1000 volts to 30,000 volts. The actual voltage depends upon the size of the CRT and its equipment application.

Voltage multipliers may also be used as primary power supplies where a 177-volt a.c. input is rectified to pulsating d.c. This d.c. output voltage may be increased (through use of a voltage multiplier) to as much as 1000 volts d.c. This voltage is generally used as the plate or screen grid voltage for electron tubes.

If you have studied transformers, you may have learned that when voltage is stepped up, the output current decreases. This is also true of voltage multipliers. Although the measured output voltage of a voltage multiplier may be several times greater than the input voltage, once a load is connected the value of the output voltage decreases. Also any small fluctuation of load impedance causes a large fluctuation in the output voltage of the multiplier. For this reason, voltage multipliers are used only in special applications where the load is constant and has a high impedance or where input voltage stability is not critical.

Voltage multipliers may be classified as voltage doublers, triplers, or quadruplers. The classification depends on the ratio of the output voltage to the input voltage. For example, a voltage multiplier that increases the peak input voltage twice is called a voltage doubler. Voltage multipliers increase voltages through the use of series-aiding voltage sources. This can be compared to the connection of dry cells (batteries) in series.

The figures used in the explanation of voltage multipliers show a transformer input, even though for some applications a transformer is not necessary. The input could be directly from the power source or line voltage. This, of course, does not isolate the equipment from the line and creates a potentially hazardous condition. Most military equipment uses transformers to minimize this hazard.

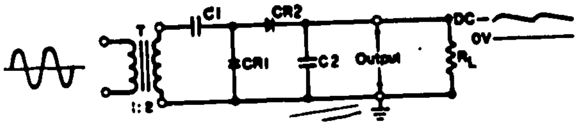
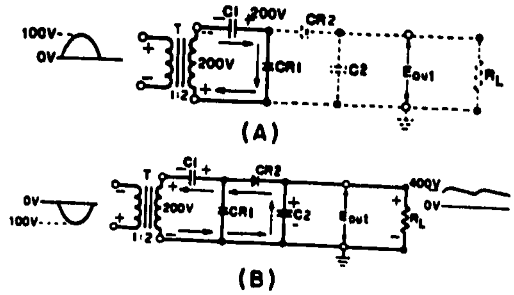


Fig 4-44. Half-wave voltage doubler.

Figure 4-44 shows the schematic for a half-wave voltage doubler. Notice the similarities between this schematic and those of half-wave voltage rectifiers with which you are already familiar. In fact, the doubler shown is made up of two half-wave voltage rectifiers. C1 and CR1 make up one half-wave rectifier, and C2 and CR2 make up the other. The schematic of the first half-wave rectifier is indicated by the dark lines in view (A) of figure 4-45. The dotted lines and associated components represent the other half-wave rectifier and load resistor.

Notice that C1 and CR1 work exactly like a half-wave rectifier. During the positive alternation of the input cycle (view A), the polarity across the secondary winding of the transformer is as shown. Note that the top of the secondary is negative. At this time CR1 is forward biased (cathode negative in respect to the anode). This forward bias causes CR1 to function like a closed switch and allows current to flow the path indicated by the arrows. At this time, C1 charges to the peak value of the input voltage, or 200 volts, with the polarity shown.

During the period when the input cycle is negative, as shown in view (B), the polarity across the secondary of the transformer is reversed. Note specifically that the top of the secondary winding is now positive. This condition now forward biases CR2 and reverse biases CR1. A series circuit now exists consisting of C1, CR2, C2 and the secondary of the transformer. The current flow is indicated by the arrows. The secondary voltage of the transformer now aids the voltage on C1. This results in a pulsating d.c. voltage of 400 volts, as shown by the waveform. The effect of series aiding is comparable to the connection of two 200-volt batteries in series. As shown in figure 4-46, C2 charges to the sum of these voltages, or 400 volts.



A. Positive alternation
B. Negative alternation

Fig 4-45. Rectifier action of CR1 and CR2.

The schematic shown in figure 4-47 is an illustration of a half-wave voltage tripler. When you compare figures 4-46 and 4-47, you should see that the circuitry is identical except for the additional parts, components, and circuitry shown by the dotted lines. (CR3, C3 and R2 make up the additional circuitry.) By themselves, CR3, C3, and R2 make up a half-wave rectifier. Of course, if you remove the added circuitry, you will once again have a half-wave voltage doubler.

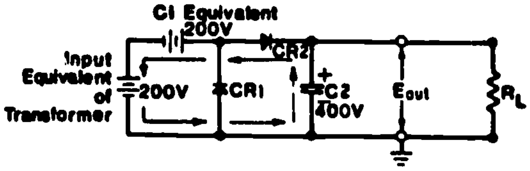


Fig 4-46. Series-aiding sources.

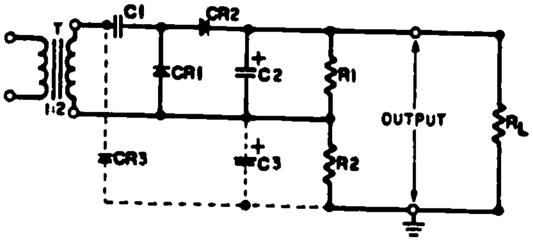
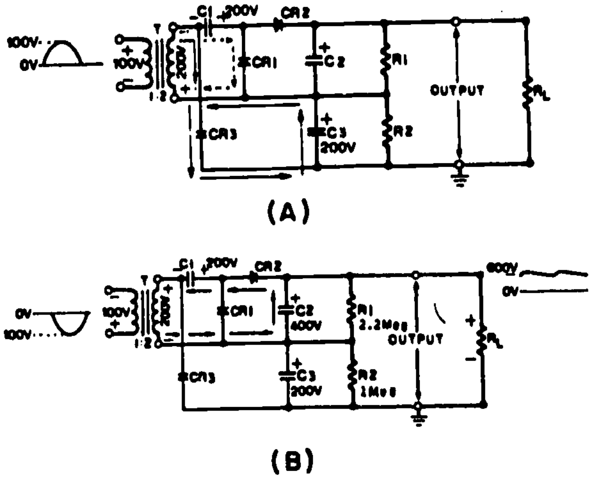


Fig 4-47. Half-wave voltage tripler.

View (A) of figure 4-48 shows the schematic for the voltage tripler. Notice that CR3 is forward biased and functions like a closed switch. This allows C3 to charge to a peak voltage of 200 volts at the same time C1 is also charging to 200 volts.

The other half of the input cycle is shown in view (B). C2 is charged to twice the input voltage, or 400 volts, as a result of the voltage-doubling action of the transformer and C1. At this time, C2 and C3 are used as series-aiding devices, and the output voltage increases to the sum of their respective voltages, or 600 volts. R1 and R2 are proportional according to the voltages across C2 and C3. In this case, there is a 2 to 1 ratio.



A. Positive alternation
B. Negative alternation

Fig 4-48. Voltage tripler.

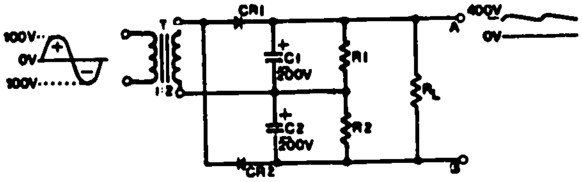


Fig 4-49. Full-wave voltage doubler.

The circuit shown in figure 4-49 is that of a full-wave voltage doubler. The main advantage of a full-wave doubler over a half-wave doubler is better voltage regulation, as a result of reduction in the output ripple amplitude and an increase in the ripple frequency. The circuit is, in fact, two half-wave rectifiers. These rectifiers function as series-aiding devices except in a slightly different way. During the alternation when the secondary of the transformer is positive at the top C1 charges to 200 volts through CR1. Then, when the transformer secondary is negative at the top, C2 charges to 200 volts through CR2. R1 and R2 are of equal value, balancing resistors which stabilize the charges of the two capacitors. Resistive load R_L is connected across C1 and C2 so that R_L receives the total charge of both capacitors. The output voltage is +400 volts when measured at the top of R_L, or point "A" with respect to point "B." If the output is measured at the bottom of R_L, it is -400 volts. Either way, the output is twice the peak value of the a.c. secondary voltage. As you may have guessed, the possibilities for voltage multiplication are almost unlimited.

Short circuit protection. The main disadvantage of a series regulator is that the pass transistor is in series with the load. If a short develops in the load, a large amount of current will flow in the regulator circuit. The pass transistor can be damaged by this excessive current flow. You can place a fuse in the circuit, but in many cases, the transistor will be damaged before the fuse blows. The best way to protect this circuit is to limit the current automatically to a safe value. A series regulator with a current-limiting circuit is shown in figure 4-50. You should recall that in order for a silicon NPN transistor to conduct, the base must be between 0.6 volt to 0.7 volt more positive than the emitter. Resistor R4 will develop a voltage drop of 0.6 volt when the load current reaches 600 milliamperes. This is illustrated using Ohm's law:

$$I = \frac{E}{R} \quad \frac{0.6 \text{ volt}}{1 \text{ ohm}} = .6 \text{ ampere or } 600 \text{ milliamperes}$$

When load current is below 600 milliamperes, the base-to-emitter voltage on Q2 is not high enough to allow Q2 to conduct. With Q2 cut off, the circuit acts like a series regulator.

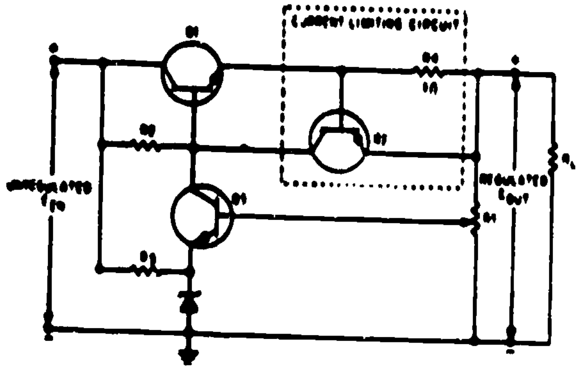


Fig 4-50. Series regulator with current limiting.

When the load current increases above 600 milliamperes, the voltage drop across R4 increases to more than 0.6 volt. This causes Q2 to conduct through resistor R2, thereby decreasing the voltage on the base of pass transistor Q1. This action causes Q1 to conduct less Therefore, the current cannot increase above 600 to 700 milliamperes.

By increasing the value of R4, you can limit the current to almost any value For example, a 100-ohm resistor develops a voltage drop of 0.6 volt at 6 milliamperes of current. You may encounter current-limiting circuits that are more sophisticated, but the theory of operation is always the same. So, if you understand this circuit, the others should be no problem.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. Circuits which maintain constant voltage or current outputs are called d.c. voltage or d.c. current ____________________.
2. The purpose of a voltage regulator is to provide an output voltage with little or no ____________________.
3. The two basic types of voltage regulators are:
   a. ____________________
   b. ____________________
4. When a series voltage regulator is used to control output voltages, any increase in the input voltage results in ____________________ in the resistance of the regulating device.
5. The shunt-type voltage regulator is connected in ____________________ with the load resistance.
6. In figure 4-37, the voltage drop across R_S and R_L determines the amount of base-emitter ____________________ for Q1.
7. In figure 4-39, view (A), when there is an increase in the input voltage, the forward bias of Q1 ____________________.
8. In view (B) of figure 4-39, when the load current increases and the output voltage momentarily drops, the resistance of Q1 ____________________ to compensate.
9. In figure 4-40, when there is an increase in the load resistance (R_L), the resistance of R_V ____________________ to compensate for the change.
10. In figure 4-43 any decrease in the base-emitter forward bias across Q1 results in ____________________ in the resistance of the transistor.
11. A half-wave voltage doubler is made up of how many half-wave rectifiers?

____

12. If a half-wave rectifier is added to a half-wave voltage doubler, the resulting circuit is a voltage ________________________________.

13. In a full-wave voltage doubler, are the capacitors connected in series or in parallel with the output load?

____

Work Unit 4-3. TROUBLE-SHOOTING POWER SUPPLIES

STATE SAFETY PRECAUTIONS WHEN WORKING WITH ELECTRONIC POWER SUPPLIES.

NAME SEVERAL FAULTY COMPONENTS THAT CAN BE IDENTIFIED THROUGH VISUAL CHECKS.

SPECIFY PROBLEMS WITHIN SPECIFIC AREAS OF A POWER SUPPLY BY USING A LOGICAL ISOLATION METHOD OF TROUBLESHOOTING.

Whenever you are working with electricity, the proper use of safety precautions is of the utmost importance to remember. In the front of all electronic technical manuals, you will always find a section on safety precautions. Also posted on each piece of equipment should be a sign listing the specific precautions for that equipment. One area that is sometimes overlooked and is a hazard is the method in which equipment is grounded. By grounding the return side of the power transformer to the metal chassis, the load being supplied by the power supply can be wired directly to the metal chassis. Thereby, the necessity of wiring directly to the return side of the transformer is eliminated. This method saves wire and reduces the cost of building the equipment. But, it solves one of the problems of the manufacturer, it creates a problem for you, the technician. Before starting to work on any electronic or electrical equipment, ALWAYS ENSURE THAT THE EQUIPMENT AND ANY TEST EQUIPMENT YOU ARE USING IS PROPERLY GROUNDED AND THAT THE RUBBER MAT YOU ARE STANDING ON IS IN GOOD CONDITION. As long as you follow these simple rules, you should be able to avoid the possibility of becoming an electrical conductor.

TESTING

There are two widely used checks in testing electronic equipment, VISUAL and SIGNAL TRACING. The importance of the visual check should not be underestimated because many technicians find defects right away simply by looking for them. A visual check does not take long; in fact, you should be able to see the problem in about two minutes if it is the kind of problem that can be seen. You should learn the following procedure. You will find yourself using it quite often as it is good not only for power supplies but also for any type of electronic equipment you may be troubleshooting. (Because diode and transistor testing were discussed earlier, it will not be discussed at this time.)

a. BEFORE YOU PLUG IN THE EQUIPMENT, LOOK FOR:

(1) Shorts. Any terminal or connection that is close to the chassis or to any other terminal should be examined for the possibility of a short. A short in any part of the power supply can cause considerable damage. Look for and remove any stray drops of solder, bits of wire, nuts, or screws. It sometimes helps to shake the chassis and listen for any tell-tale rattles. Remember to correct any problem that may cause a short circuit; if it is not causing trouble now, it may cause problems in the future.

(2) Discolored or leaking transformer. This is a sure sign that there is a short somewhere. Locate it. If the equipment has a fuse, find out why the fuse did not blow; too large a size may have been installed, or there may be a short across the fuse holder.

(3) Loose, broken, or corroded connections. Any connection that is not in good condition is a trouble spot. If it is not causing trouble now, it will probably cause problems in the future. Fix it.

(4) Damaged resistors or capacitors. A resistor that is discolored or charred has been subjected to an overload. An electrolytic capacitor will show a whitish deposit at the seal around the terminals. Check for a short whenever you notice a damaged resistor or a damaged capacitor. If there is no short, the trouble may be that the power supply has been overloaded in some way. Make a note to replace the part after signal tracing. There is no sense in risking a new part until the trouble has been located.

b. PLUG IN THE POWER SUPPLY AND LOOK FOR:

(1) Smoking parts. If any part smokes or if you hear any boiling or sputtering sounds, pull the plug immediately. There is a short circuit somewhere that you have missed in your first inspection. Use an ohmmeter to check the part once again. Start in the neighborhood of the smoking part.

(2) Sparking. Tap or shake the chassis. If you see or hear sparking, you have located a loose connection or a short. Check and repair.

If you locate and repair any of the defects listed under the visual check, make a note of what you find and what you do to correct it. It is quite probable you have found the trouble. However, a good technician takes nothing for granted. You must prove to yourself that the equipment is operating properly and that no other trouble exists.

If you find none of the defects listed under the visual check, go ahead with the SIGNAL TRACING procedure. The trouble is probably of such a nature that it cannot be seen directly with your eye - it can only be seen through the oscilloscope.

Tracing the a.c. signal through the equipment is the most rapid and accurate method of locating a trouble that cannot be found by a visual check, and it also serves as a check on any repairs you may have made. The idea is to trace the a.c. voltage from the transformer, to see it change to pulsating d.c. at the rectifier output, and then to see the pulsations smoothed out by the filter. The point where the signal stops or becomes distorted is the place to look for the trouble. If you have no d.c. output voltage, you should look for an open or a short in your signal tracing. If you have a low d.c. voltage, you should look for a defective part and keep your eyes open for the place where the signal becomes distorted.

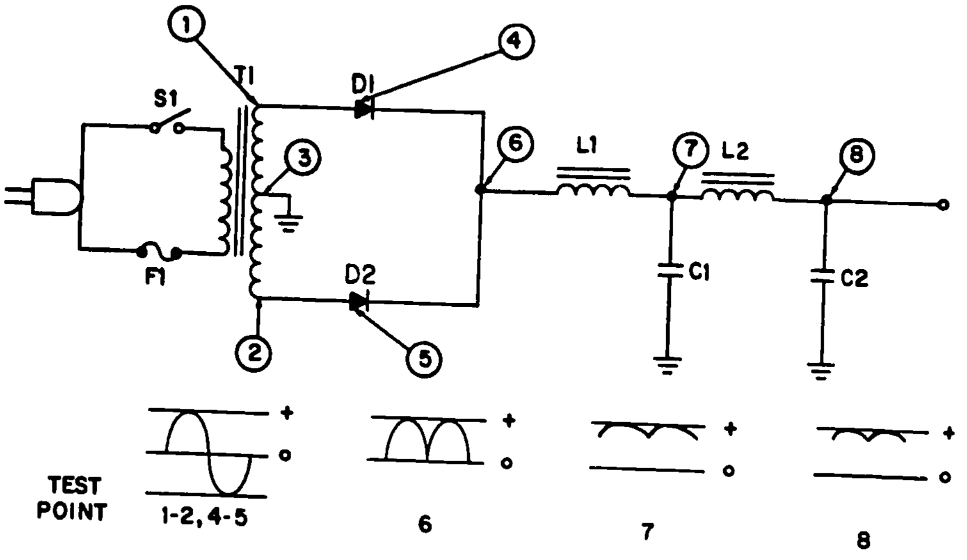


Fig 4-51. Complete power supply (without regulator).

Signal tracing is one method used to localize trouble in a circuit. This is done by observing the waveform at the input and output of each part of a circuit.

Let's review what each part of a good power supply does to a signal, as shown in figure 4-51. The a.c. voltage is brought in from the power line by means of the line cord. This voltage is connected to the primary of the transformer through the ON-OFF switch (S1). At the secondary winding of the transformer (points 1 and 2), the scope shows you a picture of the stepped-up voltage developed across each half of the secondary winding - the picture is that of a complete sine wave. Each of the two stepped-up voltages is connected between ground and one of the two anodes of the rectifier diodes. At the two rectifier anodes (points 4 and 5), there is still no change in the shape of the stepped-up voltage - the scope picture still shows a complete sine wave.

However, when you look at the scope pattern for point 6 (the voltage at the rectifier cathodes), you see the wave shape for pulsating direct current. The pulsating d.c. is fed through the first choke (L1) and filter capacitor (C1) which remove a large part of the ripple, or "hum," as shown by the wave form for point 7. Finally, the d.c. voltage is fed through the second choke (L2) and filter capacitor (C2) which remove nearly all of the remaining ripple. (See the wave form for point 8, which shows almost no visible ripple.) You now have almost pure d.c.

No matter what power supplies you use in the future, they all do the same thing - they change a.c. voltage into d.c. voltage.

The following paragraphs will give you an indication of troubles that occur with many different electronic circuit components.

As you should know by now, the transformer and the choke are quite similar in construction. Likewise, the basic troubles that they can develop are the same.

1. A winding can open.
2. Two or more turns of one winding can short together.
3. A winding can short to the casing which is usually grounded.
4. Two windings (primary and secondary) can short together (fig 4-52). This trouble is possible, of course, only in transformers.

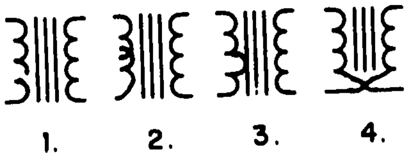


Fig 4-52. Shorted windings.

When you have decided which of these four possible troubles could be causing the symptoms, you have definite steps to take. If you surmise that there is an open winding, or windings shorted together or to ground, an ohmmeter continuity check will locate the trouble. If the turns of a winding are shorted together, you may not be able to detect a difference in winding resistance. Therefore, you need to connect a good transformer in the place of the old one and see if the symptoms are eliminated; but keep in mind that transformers are difficult to replace. Make absolutely sure that the trouble is not elsewhere in the circuit before you change the transformer.

Occasionally, the shorts will only appear when the operating voltages are applied to the transformer. In this case you might find the trouble with the megger - an instrument which applies a high voltage as it reads resistance.

Capacitor and resistor troubles. Just two things can happen to a capacitor (fig 4-53):

1. It may open up, removing the capacitor completely from the circuit.
2. It may develop an internal short circuit. This means that it begins to pass current as though it were a resistor or a direct short.

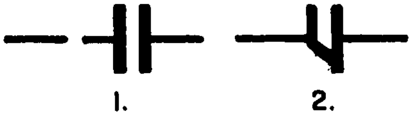


Fig 4-53. Capacitor troubles.

You may check a capacitor suspected of being open by disconnecting it from the circuit and checking it with a capacitor analyzer. You can check a capacitor suspected of being leaky with an ohmmeter; if it reads less than 500 kilohms, it is more than likely bad. However, capacitor troubles are difficult to find since they may appear intermittently or only under operating voltages. Therefore, the best check for a faulty capacitor is to replace it with one known to be good. If this restores proper operation, the fault was in the capacitor.

Resistor troubles are the simplest, but like the rest, you must keep them in mind.

1. A resistor can open.
2. A resistor can increase in value.
3. A resistor can decrease in value.

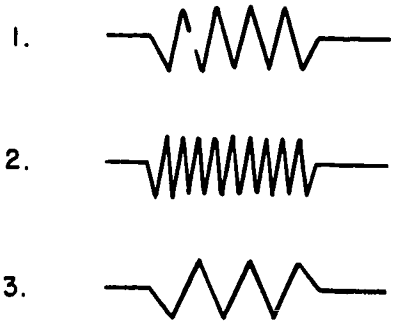


Fig 4-54. Resistor troubles.

You already know how to check possible resistor troubles. Just use an ohmmeter after making sure no parallel circuit is connected across the resistor you wish to measure. When you know a parallel circuit is connected across the resistor or when you are in doubt, disconnect one end of the resistor before measuring it. The ohmmeter check will usually be adequate. However, never forget that occasionally intermittent troubles may develop in resistors as well as in any other electronic parts.

Although you may observe problems that have not been covered specifically in this topic, you should have gained enough knowledge to localize and repair any problem that may occur.

EXERCISE: Answer the following questions and check your responses against those listed at the end of this study unit.

1. What is the most important thing to remember when troubleshooting?

   ____

2. What is the main reason for grounding the return side of the transformer to the chassis?

   ____

3. What are two types of checks used in trouble-shooting power supplies?

   a. ____________________

   b. ____________________

4. What does a discolored or leaking transformer indicate?

   ____

5. A resistor that is discolored or charred has been subjected to an ____________ ____________________.

6. You may check a capacitor suspected of being open by ______________________________

____.

7. You may check a capacitor suspected of being leaky with an ohmmeter. If it reads less than ______________, it is more than likely bad.

SUMMARY REVIEW

After completing this study unit, you have learned to identify the various components that comprise a power supply and the purpose of each component. You have learned to identify the various types of rectifier circuits and filter circuits used in a power supply. You have learned how a.c. and d.c. current flows in a power supply. Lastly, you learned how to identify faulty components through visual checks and problems within specific areas of a power supply by using a logical isolation method of trouble-shooting.

Answers to Study Unit #4 Exercises

Work Unit 4-1.

1. a. Transformer
   b. Rectifier
   c. Filter
   d. Regulator
2. To change a.c. to pulsating d.c.
3. To change pulsating d.c. to pure d.c.
4. To maintain a constant voltage to the load
5. The half-wave rectifier
6. 15.9 volts
7. It isolates the chassis from the power line.
8. The fact that the full-wave rectifier uses the full output of the transformer.
9. 120 hertz
10. 63.7 volts
11. Peak voltage is half that of the half-wave rectifier
12. The bridge rectifier can produce twice the voltage with the same size transformer as a conventional full-wave rectifier.
13. It will decrease because capacitance is inversely proportional to X_C ($X_C = \frac{1}{2\ fC}$).
14. The capacitor filter
15. In parallel
16. At a high frequency
17. A filter circuit increases the average output voltage.
18. Value of capacitance and load resistance
19. Good filtering
20. Yes
21. The cemf of the inductor
22. From 1 to 20 henries
23. Decrease
24. Increase
25. When ripple must be held at an absolute minimum
26. LC capacitor-input filter
27. a. cost
    b. size of the inductor

Work Unit 4-2.

1. regulators
2. variation
3. a. Series
   b. Shunt
4. an increase
5. in parallel
6. bias
7. increases
8. increases
9. decreases
10. an increase
11. two
12. tripler
13. in parallel

Work Unit 4-3.

1. Safety precautions
2. To eliminate shock hazard
3. a. Visual
   b. Signal tracing
4. That a short exists somewhere
5. Overload
6. disconnecting it form the circuit and checking it with a capacitor analyzer
7. 500 kilohms

SOLID STATE DEVICES

REVIEW LESSON

INSTRUCTIONS: This review lesson is designed to aid you in preparing for your final exam. You should try to complete this lesson without the aid of reference materials, but if you do not know an answer, look it up and remember what it is. The enclosed answer sheet must be filled out according to the instructions on its reverse side and mailed to MCI using the envelope provided. The questions you miss will be listed with references on a feedback sheet (MCI-R69) which will be mailed to your commanding officer with your final exam. You should study the reference material for the questions you missed before taking the final exam.

A. Multiple Choice: Select the ONE answer that BEST completes the statement or answers the question. After the corresponding number on the answer sheet, blacken the appropriate circle.

Value: 1 point each

1. Identify a solid state device?
   - a. A device that operates by magnetic motion
   - b. A device that is classified as an insulator
   - c. An electronic device which operates by virtue of the movement of electrons within a solid piece of semiconductor material
   - d. An electronic device which operates by the movement of protons

2. Identify the term negative temperature coefficient.
   - a. It is the decrease in a semiconductor's resistance as temperature rises.
   - b. It is the difference between negative and positive.
   - c. It is pure resistance.
   - d. It is the increase in a conductor's resistance as temperatures rise.

3. Space systems and computers are two of the three largest users of semiconductor devices. Identify the other large user.
   - a. Cars
   - b. Data processing equipment
   - c. Home computers
   - d. High schools

4. Identify one requirement of an electron tube which does not exist for semiconductors.
   - a. The electron tube requires a transistor to function.
   - b. The electron tube requires filament or heater voltage whereas the semiconductor device does not.
   - c. The electron tube requires a diode.
   - d. The semiconductor requires filament or heater voltage; whereas, the electron tube does not.

5. What makes the electron tube less efficient than the semiconductor?
   - a. No power input is spent by the semiconductor for conduction
   - b. The tube is smaller that the semiconductor.
   - c. No power input is spent by the electron tube for conduction.
   - d. Lower resistance exists in the electron tube.

6. Matter is defined as
   - a. anything that can not be seen.
   - b. anything that you can touch.
   - c. all liquids.
   - d. anything that occupies space and has weight.

7. Matter is made up of what three different states?
   - a. Gas, liquid, and solid
   - b. Air, water, and earth
   - c. Temperature, humidity, and solid
   - d. Temperature, moisture, and solid

8. The outer shell of an atom is called the

a. valence shell.
b. outer shell.
c. inner shell.
d. forbidden shell.

9. What determines, in terms of energy bands, whether a substance is a good insulator, semiconductor, or conductor?

a. The width of the valence shell
b. The width of the forbidden band
c. The type of material
d. The amount of voltage

10. What term describes the sharing of valence electrons between two or more atoms?

a. Valence bonding
b. Active bonding
c. Covalent bonding
d. Inactive bonding

11. Identify the two types of current flow in a semiconductor.

a. Electron and hole flow
b. Electron and shell flow
c. Shell and valence flow
d. Hole and valence flow

12. What is the name given to a doped germanium crystal with an excess of free holes?

a. N-type crystal
b. Low crystal
c. Active crystal
d. P-type crystal

13. Identify the purpose of a PN junction diode.

a. To convert alternating current into direct current
b. To activate direct current
c. To convert direct current into pulsating direct current
d. To convert pulsating direct current into voltage

14. Identify the direction of electron flow in reference to the schematic symbol for a diode.

a. Toward the arrow
b. Away from the arrow
c. With the arrow
d. Both away and with the arrows

15. In order to reverse bias in a PN junction, what terminal of a battery is connected to the P material?

a. Negative and positive terminals
b. Negative
c. Positive terminal
d. Neither terminal

16. What type of bias opposes the PN junction barrier?

a. Reverse
b. Barrier
c. Opposite
d. Forward

17. What is a load?

a. Any device that draws current from a power source
b. A device that produces voltage
c. A generator
d. A magnet

18. Identify the output of a half-wave rectifier.

a. A pulsating a.c. voltage
b. A constant d.c. voltage
c. A pulsating d.c. voltage
d. A pulsating a.c. and a constant a.c. voltage

19. Identify the type of bias that makes a diode act as a closed switch.

a. Reverse bias
b. Neutral bias
c. Forward bias
d. Negative bias

20. What type of rectifier is constructed by sandwiching a section of semiconductor material between two metal plates?

a. Metallic rectifier
b. Plate rectifier
c. Sandwiching rectifier
d. Dual rectifier

21. What is used to show how diode parameters vary over a full operating range?

a. The diode markings
b. A characteristic curve
c. A routing sheet
d. A scale

22. What is meant by diode rating?

a. The limiting values of operating conditions outside which operations could cause diode damage
b. The number that can be used in a circuit
c. The purpose for the device
d. Diodes do not have ratings.

23. What does the letter "N" indicate in the semiconductor identification system?

a. A semiconductor
b. Voltage potential
c. Current potential
d. It corresponds to the resistance of the device.

24. What type of diode has orange, blue, and gray bands?

a. 4N268
b. 2N176
c. 1N492
d. 1N368

25. What is the greatest threat to a diode?

a. Current
b. Voltage
c. Heat
d. Resistance

26. When checking a diode with an ohmmeter, what is indicated by two high resistant measurements?

a. That the diode is open or has a high forward resistance
b. That the diode is operating properly
c. That the diode is closed or has a high reverse resistance
d. That the reverse bias is too hot

27. What is the name given to the semiconductor device that has three or more elements?

a. Diode
b. Rectifier
c. Transistor
d. Emitter

28. Identify the electronic function that made the transistor famous.

a. Amplification
b. Voltage production
c. The amount of heat produced
d. Its appearance in the circuit

29. What was the name of the very first transistor?

a. H-contact
b. Hertz
c. TRW
d. Point-contact

30. What is one of the most important functions of any transistor manufacturing process?

a. The numbers of transistors produced
b. Proper identification of the device
c. Quality control
d. Doping levels

31. To properly bias an NPN transistor, what polarity voltage is applied to the collector?

a. Negative
b. Positive
c. Both negative and positive
d. High polarity voltage

32. What percentage of current in an NPN transistor reaches the collector?

a. 98 percent
b. 95 percent
c. 90 percent
d. 75 percent

33. What is the relationship to the base voltage of a properly bias NPN transistor in regards to the polarity voltage applied to the collector?

a. Equal
c. Less positive
c. More positive
d. More negative

34. What is the relationship between the polarity of the voltage applied to the PNP transistor and that applied to the NPN transistor?

a. The polarity of voltage applied to the PNP transistor is opposite of that applied to the NPN transistor.
b. There is no permanent relationship.
c. They both use the equal polarity voltage.
d. The polarity of voltage applied to the PNP transistor is the same as that applied to the NPN transistor.

35. Identify the two current loops in a transistor.

a. Emitter and polarity loop currents
b. Base and emitter loop currents
c. Base current and collector current loop
d. Forward and reverse loop currents

36. Identify the letter designation for base current.

a. C_B
b. B_C
c. I_C
d. I_B

37. What is the name of the device that provides an increase in current, voltage, or power of a signal without appreciably altering the original signal?

a. Rectifier
b. Divider
c. Amplifier
d. Emitter

38. Besides eliminating the emitter-base battery and compensating for slight variations in transistor characteristics, what other advantage can different biasing methods offer?

a. Changes in transistor conduction due to polarity
b. Changes in transistor conduction due to temperature variations
c. There are no other advantages
d. A potential balancing of the transistor

39. Identify the primary difference between the NPN and PNP amplifiers.

a. The number used in the amplifier
b. The name of the amplifier
c. The polarity of the source voltage
d. The amount of current required

40. Which of the following biasing methods is the most unstable?

a. Base current bias or fixed bias
b. Collector current bias
c. Emitter current bias
d. Element bias

41. Identify the most widely used combination-bias system.

a. The voltage-divider type
b. Zener diode type
c. High and low type
d. Point-contact type

42. What amplifier class of operation allows collector current to flow during the complete cycle of the input?

a. Class B
b. Class A
c. Class E
d. Class F

43. The two primary items that determine the class of operation of an amplifier are the amount of bias and the

a. amplitude of the input signal.
b. P-type material.
c. N-type material.
d amplitude of the output signal.

44. Which transistor configuration provides a phase reversal between the input and output signals?

a. Common base
b. Common collector
c. Common emitter
d. Common rectifier

45. Which transistor configuration has a current gain of less than 1?

a. Gamma
b. Beta
c. Charlie
d. Alpha

46. The current gain in a common-base circuit is called

a. emitter current (I_E).
b. base current (I_B).
c. collector current (I_C).
d. source current (I_S).

47. What is the formula used to compute Gamma?

a. gamma = $I_E \times I_B$
b. gamma = I_E plus I_B
c. gamma = I_E divided by I_C
d. gamma = I_E divided by I_B

48. The information normally included in the general description section of a specification sheet for a transistor is the

a. length, width, and height.
b. kind of transistor, the transistor's common applications, and mechanical data.
c. total to be used, the type, and mechanical data.
d. type installation used, application, and current rating.

49. What does the number "2" prior to the letter "N" indicate in the JAN marking scheme?

a. The number required to operate properly
b. The number of heaters in the device
c. The number of junctions in the device which in this case indicates a transistor
d. That the current reading must be divided by 2 to obtain the voltage reading

50. Identify the greatest danger to a transistor.

a. Heat
b. Cold
c. Current
d. Voltage

51. Identify the method for checking a transistor which is cumbersome when more than one transistor is bad in a circuit?

a. Forward method
b. Avalence method
c. Reverse method
d. Substitution method

52. Identify the safety precaution which must be taken before replacing a transistor.

a. Add heat to aid the operation.
b. The circuit must be energized.
c. Keep the circuit doping down.
d. The power must be removed from the circuit.

53. How is the collector lead identified on an oval-shaped transistor?

a. By the wide space between the collector lead and the other two leads (emitter and base)
b. By the color markings on the lead
c. The collector lead is shorter than the other two leads.
d. The collector lead is always on the left.

54. Identify the two most important parameters used for testing a transistor.

a. Current and resistance
b. Gain and leakage
c. Gain and resistance
d. Voltage and gain

55. Integrated circuits are defined as

a. two circuit added by conductors.
b. circuits packaged as a single circuit
c. a master circuit aided by two minor circuits.
d. two circuits separated by an insulator.

56. In a reverse-biased PN junction, which current carriers cause leakage current?

a. The majority carriers
b. The forward carriers
c. The reverse carriers
d. The minority carriers

57. What is the doping level of an avalanche effect diode when compared to the doping level of a Zener effect diode?

a. The doping level of an avalence effect diode is lower.
b. The doping levels are equal.
c. The doping level of an avalence effect diode is higher.
d. The doping level is twice that of the Zener effect diode.

58. Why is electron flow with the arrow in the symbol of a Zener diode instead of against the arrow as it is in a normal diode?

a. Because it indicates a combination bias method
b. Because Zener diodes are operated in the reverse-bias mode
c. Because that is the way all diodes are manufactured
d. Because Zener diodes are operated in the forward bias mode

59. Identify the main difference in construction between normal PN junction diodes and tunnel diodes.

a. The amount of voltage
b. The number of outer shells
c. The number of inner shells
d. The amount of doping

60. What resistance property is found in tunnel diodes but not in normal diodes?

a. Positive resistance
b. Negative resistance
c. Positive current
d. Negative current

61. In the tunnel diode, the tunneling current is at what level when the forbidden gap of the N-type material is at the same energy level as the empty states of the P-type material?

a. Maximum
b. Minimum
c. The same level
d. Doping level

62. The varactor displays what useful electrical property?

a. Variable impedance
b. Limiting
c. Variable capacitance
d. Active

63. When the reverse bias on a varactor is increased, what, if anything, happens to the effective capacitance?

a. The capacitance decreases
b. The capacitance increases
c. Nothing happens
d. The capacitance will be balanced

64. The SCR is primarily used for what function?

a. To switching power on or off
b. To switch power on
c. To switch power off
d. As a filter

65. When an SCR is forward biased, what is needed to cause it to conduct?

a. The zener signal
b. Collector current
c. The cathode signal
d. The gate signal

66. What is the only way to cause an SCR to stop conducting?

a. By reducing reverse bias
b. By increasing the doping level
c. By decreasing the doping level
d. The forward bias must be reduced below the minimum conduction level

67. The triac is similar in operation to what device?

a. SCR
b. Transistor
c. Diode
d. LED

68. When used for a.c. current control, during which alternation of the a.c. cycle does the triac control current flow?

a. During both alternations
b. During the positive alternation
c. During the negative alternation
d. During neither alternation

69. The functions of a triac can be duplicated by connecting two actual

a. diodes.
b. transistors.
c. SCRs.
d. Zener diodes.

70. The LED, photocell, and solar cell are three of the five types of optoelectronic devices. Identify the remaining two.

a. Transistor and zener diode
b. Photodiode and the phototransistor
c. Thyratron and SCR
d. SCR and triac

71. Identify the type of bias that is required to cause an LED to produce light.

a. Reverse bias
b. Both reverse and forward bias
c. Neutral bias
d. Forward bias

72. When compared to incandescent lamps, what is the power requirement of an LED?

a. Very low
b. Very high
c. Equal
d. Zero

73. In a common anode, seven-segment LED display, an individual LED will light if a negative voltage is applied to what element?

a. The cathode
b. The anode
c. The base
d. The point emitter

74. What is the resistance level of a photodiode in total darkness?

a. Equal
b. Zero
c. Very low
d. Very high

75. What is a typical light-to-darkness resistance ratio for a photocell?

a. 2:1
b. 1:500
c. 1:1000
d. 2:1000

76. Identify the semiconductor device that produces electrical energy when exposed to light?

a. Photovoltaic cell
b. Transistor
c. PhotoLED
d. Diode

77. The unijunction transistor (UJT) has how many PN junctions?

a. Three
b. Two
c. One
d. Zero

78. The area between base 1 and base 2 in a unijunction transistor acts as what type of common circuit component?

a. Variable resistor
b. SCR
c. Triac
d. Voltage divider

79. Identify the name of the sequential rise in voltage between the two bases of the unijunction transistor.

a. Gage bias
b. Voltage bias
c. A gate gradient
d. A voltage gradient

80. Identify the normal current path for a unijunction transistor.

a. From base 2 to the emitter
b. From base 2 to the collector
c. From base 1 to the emitter
d. From base 1 to the collector

81. Identify the one advantage of the field effect transistor (FET) when compared to the bipolar transistor.

a. High input capacitance
b. Low input impedance
c. High input impedance
d. Low input capacitance

82. Identify two types of junction field effect transistor?

a. Forward and reverse
b. N-channel and P-channel
c. Voltage and current
d. Resistance and current

83. Identify the key to the field effect transistor (FET) operation.

a. Effective cross-sectional area of the channel
b. The amount of current flow
c. A positive gate voltage
d. Reverse biasing

84. When compared to the junction field effect transistor (JFET), what is the input impedance of the metal oxide semiconductor field effect transistor (MOSFET)?

a. The MOSFET has a lower input impedance
b. There is no difference.
c. It is balanced.
d. The MOSFET has a higher input impedance.

85. What are the four elements of the MOSFET?

a. Gate, heater, filament, and drain
b. Gate, source, drain, and substrate
c. Source, gate, heater, and filament
d. Drain, resistance, source, and substrate

86. The MOSFET is normally constructed so that it can operate in one of two basic modes; the depletion mode or the

a. negative mode.
b. filter mode.
c. enhancement mode.
d. regulator mode.

87. Identify the four sections of a power supply.

a. Rectifier, diode, gate, and filter
b. Transformer, rectifier, filter, regulator
c. Filter, regulator, switch, and gate
d. Regulator, filter, switch, and source

88. Identify the section of the power supply that changes a.c. to pulsating d.c.

a. Rectifier section
b. Gate section
c. Filter section
d. Diode section

89. Identify the section of the power supply that changes pulsating d.c. to pure d.c.

a. Gate
b. Transformer
c. Rectifier
d. Filter section

90. What is the purpose of the regulator section of the power supply?

a. To maintain a constant voltage to the load
b. To decrease the voltage to the load
c. To separate the output of the source
d. To maintain a constant sinewave

91. In addition to stepping up or stepping down the input line voltage, what additional purpose does the transformer serve?

a. To transform a.c. to d.c.
b. It changes current direction
c. It isolates the chassis from the power line
d. None

92. What is the major factor that led to the development of the full-wave rectifier?

a. The fact that the full-wave rectifier uses the full output of the transformer
b. It keeps the LEDS operating properly.
c. It uses fewer SCRs.
d. The fact that it uses the full ouput of the transformer

93. What is the average voltage (E_{avg}) output of a full-wave rectifier with an output of 100 volts peak?

a. 53.7 volts
b. 60.7 volts
c. 63.7 volts
d. 67.3 volts

94. What is the main disadvantage of a conventional full-wave rectifier?

a. Peak current is twice that of the half-wave rectifier
b. Peak voltage is twice that of the half-wave rectifier
c. Peak current is half that of the half-wave rectifier
d. Peak voltage is half that of the half-wave rectifier

95. What main advantage does a bridge rectifier have over a conventional full-wave rectifier?

a. Current limiting ability
b. It is better insulated for your protection
c. The transformer gets hot.
d. The bridge rectifier can produce twice the voltage with the same size transformer.

96. Identify the most basic type of filter.

a. The capacitor filter
b. The resistance filter
c. The inductance filter
d. The voltage filter

97. The rate of discharge of a capacitor in a filter circuit is determined by the

a. value of capacitance and load resistance.
b. size of the circuit.
c. gate signal.
d. circuit design.

98. Identify what prevents the rapid charging of the capacitor in an LC choke-input filter?

a. The ripple frequency
b. Change time
c. Pulsating voltage
d. The cemf of the inductor

99. If the impedance of the choke is increased, what will happen to the ripple frequency?

a. It will decrease
b. It will increase
c. It will remain the same.
d. It will become unequal.

100. When is a second RC filter stage used?

a. When the ripple must be held at an absolute maximum
b. When the pulsating voltage is first applied
c. When the ripple must be held at an absolute minimum
d. When the capacitor is connected across the output

101. Circuits which maintain constant voltage or current outputs are called d.c. voltage or d.c. current

a. filters.
b. rectifiers.
c. regulators.
d. transformers.

102. The purpose of a voltage regulator is to provide an output voltage with little or no

a. resistance.
b. current.
c. impedance.
d. variation.

103. Identify the two basic types of voltage regulators.

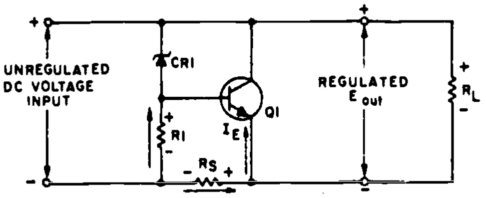
a. Half-wave and full-wave
b. Series and shunt
c. Parallel and binder
d. Series and binder

104. When a series voltage regulator is used to control output voltages, any increase in the input voltages results in _______________________ in the resistance of the regulating device.

a. a decrease
b. a ripple effect
c. an increase
d. filtering action

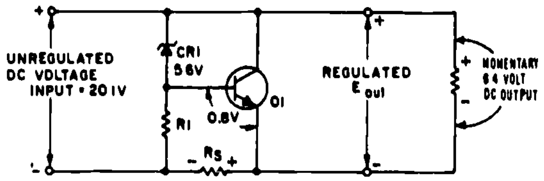
105. Identify how a shunt-type voltage regulator is connected with the load resistance.

a. In parallel
b. In series
c. In series-parallel
d. Positive in respect to the cathode



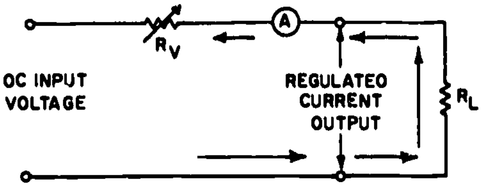
106. In the figure above, the voltage drop across R_S and R1 determines the amount of base-emitter __________________ for Q1.

a. bias
b. resistance
c. filtering
d. ripple frequency



107. In the figure above, when the load current increases and the output voltage momentarily drops, the resistance of Q1 ______________ to compensate.

a. decreases
b. remains constant
c. reduces to half
d. increases



108. In the figure above, when there is an increase in the load resistance (R_L), t resistance of R_V ____________________ to compensate for the change.

a. decreases
b. remains constant
c. reduces to one-tenth
d. increases

109. A half-wave voltage doubler is made up of how many half-wave rectifiers?

a. Four
b. Three
c. Two
d. One

110. If a half-wave rectifier is added to a half-wave voltage doubler, the resulting circuit is a voltage

a. doubler.
b. tripler.
c. rectifier.
d. triac.

111. In a full-wave voltage divider, how are the capacitors connected with the output load?

a. In series
b. In parallel
c. In parallel-series combination
d. They are not connected.

112. Identify the most important thing to remember when trouble-shooting.

a. The type circuit
b. The total voltage
c. The type equipment being used
d. Safety precautions

113. What is the main reason for grounding the return side of the transformer to the chassis?

a. To prevent damage to the transformer
b. For the proper operation of the equipment
c. Because it is a Marine Corps requirement
d. To eliminate electrical shock

114. What are two types of checks used in trouble-shooting power supplies?

a. Mechanical and chemical
b. Physical and visual
c. Visual and signal tracing
d. Rogets and Thesaurus tracing

115. Identify what is indicated by a discolored or leaking transformer.

a. That a short exists somewhere
b. An open circuit
c. Significant changes in design
d. Solid state components

116. A discolored or charred resistor has been subjected to a(n)

a. overload.
b. high resistance.
c. inexpericened technician.
d. filter.

117. To check a capacitor suspected of being open, first disconnect it from the circuit and check it with a(n)

a. capacitor analyzer.
b. voltmeter.
c. ohmmeter.
d. SCR.

118. You may check a capacitor suspected of being leaky with an ohmmeter. If it reads less than ______________, it is more than likely bad.

a. 300 kilohms
b. 500 kilohms
c. 750 kilohms
d. 1000 kilohms

Total Points: 118

*U.S. GOVERNMENT PRINTING OFFICE: 1986-491-630:40037

STUDENT COURSE CONTENT ASSISTANCE REQUEST

COURSE NUMBER	COURSE TITLE	DATE:
NAME	RANK	MOS
SOCIAL SECURITY NUMBER		DATE:
COMPLETE MILITARY ADDRESS: (INCLUDE RUC)		
UNIT PHONE NUMBER		(PLEASE FILL IN ADDRESS ON REVERSE SIDE)

1. Use this form for any questions you may have about this course. Write out your question and refer to the study unit, work unit, or study question which you are having problems with. Complete the self-addressed block on the reverse side. Before mailing, fold the form and staple it so that MCI's address is showing. Additional sheets may be attached to this side of the form. Your question will be answered promptly by a Marine NCO.

MY QUESTION IS: ____________

OUR ANSWER IS: ____________

SSTD-1

DEPARTMENT OF THE NAVY

UNITED STATES MARINE CORPS
MARINE CORPS INSTITUTE
ARLINGTON, VA 22222-0001

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

NAME	RANK	SSN
MILITARY ADDRESS		

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 12486 WASH., D.C.

POSTAGE WILL BE PAID BY U.S. MARINE CORPS

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

UNITED STATES MARINE CORPS
MARINE CORPS INSTITUTE
ARLINGTON, VA 22222-0001

STUDENT REQUEST/INQUIRY
MCI - R11

DATE SENT:__________
DATE RECIEVED AT MCI:__________

COURSE NUMBER	COURSE TITLE

COMPLETE ALL PORTIONS OF SECTION 1

Section 1. Student Identification

Rank	Initials	Last Name	MOS

SSN

REPORTING UNIT CODE (RUC)

MILITARY ADDRESS

____

____________________ ZIP CODE

INSTRUCTIONS: Print or type name, rank and address clearly. Include ZIP CODE. Only Class III Reservists use civilian address

- -

Section 2. CHECK THE APPROPRIATE BOX AND FILL IN THE APPROPRIATE SPACES.

FOR REGULAR AND CLASS II RESERVE MARINES THIS FORM MUST BE SIGNED BY THE COMMANDING OFFICER OR HIS REPRESENTATIVE, i.e. TRAINING NCO.

1. EXTENSION - Please grant an extension (will not be granted if already on extension.)
2. NOTICE OF COURSE COMPLETION - Final Exam Sent On_______. (New exam will be sent if exam not received at MCI.)
3. REENROLLMENT - Student has course materials (See para. 4003 of Vol. I of MCI Catalog for information on reenrollment.)
4. OVERDUE FINAL EXAM - Last (Review) lesson sent on __________. Please send exam.
5. Please send new ANSWER SHEETS.
6. Please send missing course materials (Not included in course package.)
   Lessons____ Manual_____ Other______________
7. CHANGE - Rank_______ Name______________
   Social Security Number
   RUC___________
8. (explain)______________________
   ____

FOR MCI USE ONLY

1. D____
2. On SMF____ K____ L____
3. C____
4. L____
5. Q____
6. P____
7. E____
8. __________________
   ____
   ____

DATE COMPLETED____________

ORIGINATOR CODE____________

Note: This form will not be returned by MCI. If request is valid, transaction will show on next UAR or on MCI-R1_ form.

____
SIGNATURE-TITLE OR RANK
(MUST BE CO OR REPRESENTATIVE)

- -

STUDENT: Detach and retain this portion.

DATA REQUIRED BY THE PRIVACY ACT OF 1974
(5 U.S.C. 522A)

1. AUTHORITY: Title 5, USC, Sec. 301. Use of your Social Security Number is authorized by Executive Order 9397 of 22 Nov 43.

2. PRINCIPAL PURPOSE: The Student Course Content Assistance Request is used to transmit information concerning student participation in MCI courses.

3. ROUTINE USE: This information is used by MCI personnel to research student inquires. In some cases information contained therein is used to update correspondence course and individual student records maintained by the Marine Corps Institute.

4. MANDATORY OR VOLUNTARY DISCLOSURE AND EFFECT ON INDIVIDUAL NOT PROVIDING INFORMATION: Disclosure is voluntary. Failure to provide information may result in the provision of incomplete service to your inquiry. Failure to provide your Social Security Number will delay the processing of your assistance request.

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 12495 WASHINGTON D. C.

POSTAGE WILL BE PAID BY U.S. MARINE CORPS

UNITED STATES MARINE CORPS
MARINE CORPS INSTITUTE
ARLINGTON, VA 22222-0001

DETACH BEFORE MAILING