ACHIEVEMENT THROUGH ELECTRONICS

nri

INTRODUCING YOU TO ELECTRONICS

B101

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

INTRODUCING YOU TO ELECTRONICS

B101

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

☒ 1. Introduction .... Pages 1 - 4
This section gives you a glimpse of the opportunities in store for you, and also some helpful suggestions on how to study.

☒ 2. Electricity .... Pages 5 - 12
You get a basic idea of what electricity is.

☒ 3. Current Flow .... Pages 13 - 17
You learn how current is made to flow in a circuit, and the relationship between current, voltage and resistance.

☐ 4. Magnetism .... Pages 18 - 24
You learn about the importance of magnetism in electronics.

☐ 5. Electronic Components .... Pages 25 - 37
We look at a number of basic parts; conductors, insulators, batteries, coils, transformers, resistors, vacuum tubes and transistors.

☐ 6. Answers to Self-Test Questions .... Pages 38 - 39

☐ 7. Answer the Lesson Questions.

☐ 8. Start Studying the Next Lesson.

PII3M1174 1974 EDITION Litho in U.S.A.

INTRODUCING YOU TO ELECTRONICS

Deciding to study the NRI course is one of the wisest decisions of your life. Why? Because this course has been planned and written from beginning to end especially for men who must do their studying at home, usually after their regular day's work. Each of NRI's bite size lessons has been programmed into small sections. Most of these sections can be mastered with little over an hour's diligent study.

The amount of education you now have is of less importance than an unswerving ambition to succeed. The first lesson starts right at the beginning assuming that you know nothing about electronics, and prepares you for the more advanced second lesson. If you will make a firm resolution now to work on your course regularly, you can be assured of success. The rewards that can be yours by satisfactorily completing your course are within your grasp. You've taken the first step by enrolling, now the next step is to set up a study schedule, and then stick to that study schedule at all costs.

Since in the second book, and in each of the following books that you will study, you will be building on what you have learned in previous books, it is important for you to understand each new idea as it is presented before going on to the next subject. Do not make the mistake of skipping over a section that you do not at first understand. Usually if any section of a lesson at first seems difficult, you will be able to understand it completely by rereading it several times. If after you have carefully studied a section of a lesson, you find there are still some points puzzling you, take advantage of the NRI Consultation Service to ask for extra help. Try to be as specific as possible about exactly what is not clear to you -- we will be glad to assist you; we are here to help you.

The surest way to succeed is to be determined to succeed. Make a sincere promise to yourself right now, that you are going to complete your NRI course and succeed in electronics.

OPPORTUNITIES

Let us look at a few of the opportunities that are available to the qualified technician. As an NRI graduate you will be in a position to choose your opportunity, instead of waiting for it to come to you.

In radio, for instance, there are hundreds of different job opportunities in entertainment broadcasting alone. You might think that television has eliminated the opportunities in this field of radio. The actual facts are that almost every week new radio stations are licensed by the FCC and that the annual radio receiver production in this country is close to 10,000,000 receivers.

Television has not eliminated opportunities in the radio field; instead it has created new opportunities of its own. The demand for broadcast technicians and television servicemen has never been fully satisfied since television first swept across the country. Now color television is here, and it is creating still greater demands for trained men. You can be one of them.

Another fact that people fail to realize is that there are many opportunities outside the entertainment broadcasting field. There are thousands of fascinating well-paid jobs for men in communications systems on land, on sea, and in the air. Fire and police departments, telephone companies, power companies, gas companies, railroads, and airlines are only a few of the many organizations using radio communications equipment. The use of radio in industrial communications is in fact increasing so fast that finding frequency assignments for all the new stations is becoming a serious problem.

Electronic equipment is also becoming more and more important in industry. Electronics in industry represents a tremendous new field, the surface of which has just been scratched. Electronic equipment is used to count finished components coming off assembly lines, it is used to inspect manufactured parts, it is used to control precision machines, automatically making possible the high-speed production of items that could formerly be made only manually by highly skilled operators. Electronics is used in oil refineries, in the manufacturing and quality-control of new cars, in the new plastics industry and in many other fields, too numerous to mention. As a matter of fact, there is hardly an industrial process in which electronics cannot be put to use advantageously. Here is a field of unlimited opportunities; a field that is just developing.

From the preceding list of opportunities you might think that this is an industry already fully developed. Actually we have barely scratched the surface. You are going to see breath-taking new developments which far outshadow even the miracle of color television. Because your NRI training is built upon a sound foundation, you will be prepared for the undreamed-of jobs that will soon be created by new developments. You will not only learn about equipment in use today, but you will also learn the fundamental ideas in back of the operation of tomorrow's equipment. Once you understand this basic theory, you will be able to understand new developments as they come along. You will understand how they work, because you will know the fundamentals. New electronic equipment will not use new circuits, it will use the basic circuits you will be studying, but in new ways.

HOW TO STUDY

Naturally, you want to complete your course and become an expert technician as quickly as possible. To help you do this, here are a few suggestions on how to study.

A Study Schedule.

The first suggestion is that you plan a study schedule. Decide how many hours you'll study each week. Then decide on which days you will be able to work on your course. Finally, decide the time during the day or evening that will be best for studying.

How Often to Study.

Space your study periods close together. The ideal arrangement is to devote some time to your course each day. If you can't spare one or two hours, study for 45 minutes, half an hour, or whatever you can spare.

Regular study is the key to learning effectively. You can learn and remember much more from studying 30 minutes daily than from longer sessions spaced several days apart. Studying daily, or every other day, keeps the instruction fresh in your mind. If three or four days pass between study periods, you may forget most of what you learned from your last session. It's hard to pick up where you left off, and you often have to do a lot of "back-tracking" to refresh your memory.

Take a Break.

Unless you're accustomed to studying regularly, you may find that you tire easily. This isn't unusual. It will take time for you to get the "study habit".

In the beginning, make your study periods short and break them up with periods of rest. After you've completed several lessons, you will be able to study longer without getting tired.

Be Fair to Yourself.

Most of us have asked for that second slice of apple pie only to discover we are too full to handle it. An over-ambitious study schedule can be like that second slice of pie. Be sure your study schedule is reasonable, one that you will have no trouble following. This schedule need not deprive you of time for other activities. A well planned schedule will see to it that each activity gets its fair share of your attention.

If possible, have a regular place to study, and get used to going to the same place each time you study. Naturally, this place should be relatively free of distracting noises. If the noise around you can't be controlled, consider going to a Public Library. Libraries are quiet and have an ideal study atmosphere.

When you sit down to study, be sure to have everything you need. Then you won't have to stop in the middle of a lesson to get a pencil, paper etc. Keep these materials near your place of study at all times so that once you sit down, you can go right to work.

Here are some suggestions to help you study your lesson texts systematically. These suggestions are the result of years of experience with well over 600,000 students and will enable you to learn effectively.

Survey the Lesson.

Begin each new lesson with a survey. First, read the Study Schedule to get a general idea of the subjects and their order of presentation. Next, thumb through the lesson and look at the different main headings. The main headings are the same as each step in the Study Schedule.

Now carry the survey further. Look at the smaller headings under each main heading to get a more detailed idea of what the lesson covers. In fact, you should glance at the first two or three sentences under each heading.

The survey acquaints you with the lesson. When you begin a thorough study, you will know what to expect.

Read and Recite.

After the survey, you are ready to study the lesson. Remember, the text is full of facts and explanations. To absorb them, you should read every sentence carefully, turning over in your mind what it says and means.

Stop periodically and try to recall what you've read. Actually recite to yourself the main headings and the important ideas under each. Then check back to see whether you've covered everything. If you've left something out, restudy those points and again try to repeat them to yourself.

Answering the Test Questions.

On the last page of each lesson are ten questions covering subjects in the lesson text.

You'll probably find many of the questions easy to answer. Some will require a good deal of thought. A few may seem difficult but will help you develop the ability to work out problems you'll encounter in the field.

If you come to a question you can't answer, find that part of the lesson where the subject is covered. Reread that portion carefully, fixing the answer to the question firmly in mind. Then write out the answer in your own words.

You will find it helpful to review periodically. A good system is to review completed lessons at about the same rate you study new lessons.

For example, when you finish six lessons, go over the first two again. When you finish your next new lesson, review the third and fourth lessons. After your review catches up to your new lessons, start the review over again. Naturally, when you review, you will not have to be as thorough as when you study the lesson. However, just skimming through the lesson you are reviewing will help recall the important points in that lesson.

Your lessons contain a series of self-test questions which are related to the major topics of each text. You should try to answer each question, in writing, as you come to it in your studies. Before continuing, check your answer against the "Answers to Self-Test Questions" on pages 38 and 39. If you have answered the questions correctly, go on to the next section. If any of your answers are incorrect, review the topic just covered.

These Self-Test Questions are for your own use so you can see how well you understand the material. Do NOT Send the Answers to the Self-Test Questions to NRI for Grading. This will only slow down your instructor.

Now you are ready to go -- ready to start your study of electronics. At first we will study a few basic ideas and then we will use these ideas in building up some simple circuits. Make sure you fully understand the ideas and circuits as they are presented. Even the most complex electronic equipment is made up of nothing more than a large number of the simple circuits that you will study in your early lessons.

Electricity

Electricity and magnetism play an important part in the operation of all electronic equipment. But what is electricity? What is magnetism? If you learn the answers to these two questions once and for all now, you will have the foundation for everything else you will study and work with in your electronics career. Let us take up electricity first.

WHAT ELECTRICITY IS

We will start our discussion of what electricity is by describing a few simple experiments that were performed by scientists many years ago. These experiments helped lead to the understanding of electricity that we have today. You do not have to perform these experiments. We are describing them only to help you understand electricity.

If two small glass balls are suspended by threads as shown in Fig. 1, the weight of the glass balls will cause them to hang straight down as shown. Now if we take a piece of silk

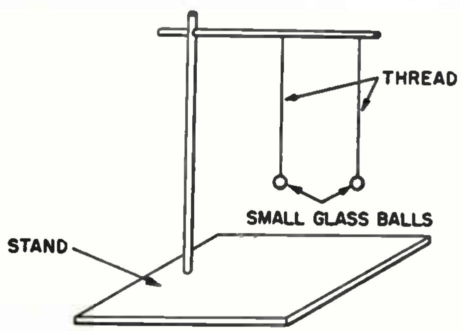


Fig. 1. Two small glass balls suspended by threads will hang vertically because of the weight of the balls.

cloth and rub the balls with the cloth and then suspend them, instead of hanging straight down as they did before, the balls will tend to swing out as shown in Fig. 2 as though some force were pushing them apart.

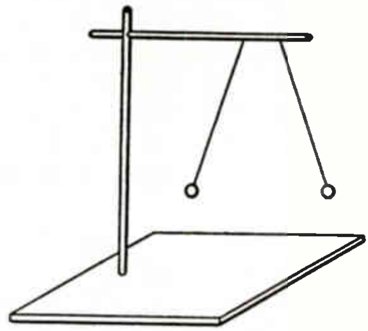


Fig. 2. Two glass balls that have been rubbed with a silk cloth move apart.

Actually, there must be some force pushing them apart; rubbing the balls with the silk cloth has produced the force.

The same experiment can be performed using two small hard-rubber balls. Again, before the balls have been rubbed with the silk cloth they will hang vertically, but once they have been rubbed with the silk cloth they will push apart. Again rubbing the small balls with the silk cloth has produced a force which pushes the balls apart.

There are two important points illustrated by these experiments. First, rubbing the balls with the silk cloth produced a force that pushed the balls apart. We call this force a "charge". The act of producing this force is called "charging". We say that the balls are "charged."

Second, let us consider what happened when we rubbed the glass balls with the silk cloth. When we rubbed the two balls, we charged them. Since both balls were charged in the same way, it is logical to assume that we have placed the same kind of a charge on the two balls. But the two balls pushed each other apart. We had similar results when we charged the two rubber balls. They also repelled each other. From these experiments we can conclude that if two objects are charged with the same kind of a charge, they will repel each other. This is a basic electrical law, and it is usually stated:

like charges repel.

Now, if the same experiment is performed using one small glass ball and one small rubber ball, we would observe an entirely different effect. When the balls are rubbed with the silk and then suspended, instead of moving apart, the two balls would move toward each other as shown in Fig. 3. If they are placed close enough together, they will move toward each other until they touch. Once the two balls touch each other, they will begin to move apart to hang straight down. They will probably swing past the straight-down point and then swing back together and touch again. This cycle may be repeated several times until eventually the balls will hang straight down as they would if they had not been charged in the first place. Now let us see what conclusions we can draw from this experiment.

We have seen that rubbing the two balls with the silk cloth charged them as before. This we know is true because a force was produced on the two balls, otherwise they would simply hang straight down. We also know that the forces produced on the two balls were such that the balls did not repel each other, at least at first, because the two balls moved together and touched. Since we have seen from the previous experiments that like charges repel, these charges must have been unlike. In other words, there must be a different kind of charge on the glass ball from the one on the rubber.

These simple experiments lead us to a fundamental important rule,

"like charges repel;
unlike charges attract."

Remember this rule, it is important. You will use it throughout your entire career in electronics. You will soon use this simple rule to help explain the operation of many electronic devices.

Now let us proceed with our study of electricity to see if we can explain more fully what happened in the experiments we have just described. To do this we must study the electron theory.

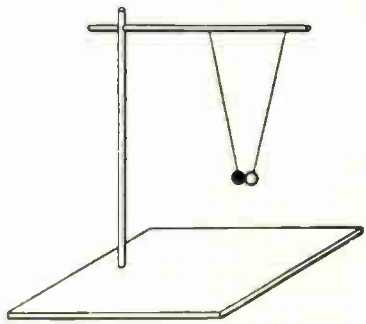


Fig. 3. The charged glass and rubber balls will move together and touch each other.

THE ELECTRON THEORY

Everything on this earth is made up of tiny particles. You can see for yourself that the earth is not one solid piece of material, it is made up of tiny particles of sand and stone and rock. Even the smallest grain of sand is itself made up of millions of still smaller particles, so small that they cannot be seen even with the most powerful microscope.

The smallest particle of a substance that retains the original properties of the substance is called an "atom." All atoms of a given substance are alike. In other words, the smallest particle of a piece of copper that still is and resembles copper is called an atom. These atoms are so small that a piece of copper the size of the head of a pin would contain millions of atoms.

But the atom is not the smallest particle, the atom itself is made up of still smaller particles. Scientists have identified a number of different particles from which the atom is made. However, we are interested in only two of these particles, the nucleus* and the electron. We are more interested in the electron than we are in the nucleus.

The nucleus is the center of the atom. Travelling around the nucleus in elliptical paths (a somewhat circular path that has been squashed, like an egg or football) will be one or more electrons. The number of electrons will be different for atoms of different elements.

The nucleus of an atom has a positive charge. The simplest atom is the hydrogen atom and the nucleus of this atom has one positive charge. Travelling around this nucleus in an elliptical path is one electron which has a negative charge. The negative charge on the electron exactly balances the positive charge on the nucleus so that electrically the atom is neutral. The most complex atom found in nature is the uranium atom. The nucleus of the uranium atom has 92 positive charges and travelling around the nucleus of this atom are 92 electrons which will exactly balance the 92 charges on the nucleus, so that the net electrical charge on the atom is zero.

Between the simplest atom, which is the hydrogen atom, and the most complex atom are 90 other materials. They range from the helium atom which has two positive charges on the nucleus and two electrons travelling around it up to the protactinium atom which has 91 charges on the nucleus and 91 electrons travelling around it.

In their natural state the electrons travelling around the nucleus of an atom exactly balance the positive charges on the nucleus. However, under some circumstances an atom might loose one of its electrons. When this happens, the electron, which carries a negative electrical charge, moves off into space or over to a nearby atom. Meanwhile, the atom which has lost the electron now does not have enough electrons to completely balance the positive charge on the nucleus. As a result, the atom has a positive charge.

Under some circumstances the opposite might happen, and an atom might pick up an extra electron.

* To be strictly correct we should not call the nucleus a particle, because it is made up of smaller particles. However, for our purposes we can consider the nucleus as one particle.

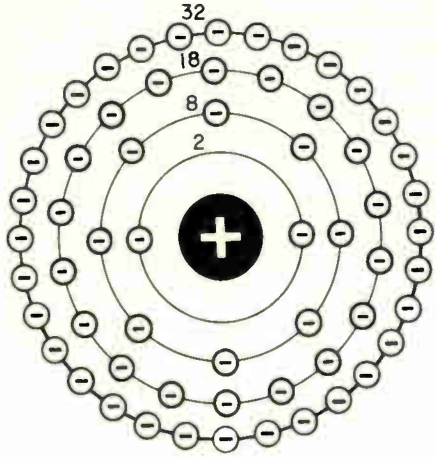


Fig. 4. The maximum number of electrons there can be in each of the first four rings of any atom.

When this happens the atom has more electrons than it needs to completely neutralize the charge on the nucleus. As a result, the atom will have a negative charge.

The electrons around the nucleus of an atom travel around it in rings. There is a maximum number of electrons that can be in each ring. In Fig. 4 we have shown the maximum number of electrons that can be in each of the first four rings surrounding the nucleus of an atom. As the atoms go from hydrogen, the simplest atom, to uranium, the most complex atom, the electrons fill the inner rings first. For example, the hydrogen atom shown in Fig. 5A has one electron traveling around the nucleus. The helium atom shown in Fig. 5B has the first ring filled with two electrons travelling around the nucleus. The lithium atom has three electrons as in Fig. 5C, two electrons fill the first ring and the third electron appears in the second ring. Subsequent elements will have electrons in the second ring until a maximum of eight electrons is reached in this ring. The next element will have two electrons in the first ring, eight in the second and the eleventh electron in the third ring.

When the outer ring of electrons in an atom is filled, the atom is very stable electrically and chemically. It is almost impossible to get an electron to move out of the atom or to force another electron into the atom. On the other hand, if the outer ring has all the electrons it can hold except one, it is very easy to force an electron into that outer ring to fill the ring. By contrast, if the outer ring has only one electron in it, that electron is not held very closely to the atom and therefore it can easily move out of its position into space

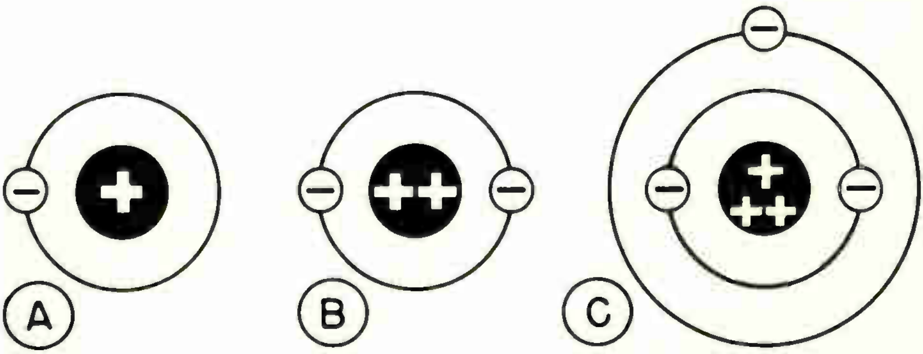


Fig. 5. The hydrogen atom is shown at A, the helium atom at B, and the lithium atom at C.

or to another atom.

The copper atom is an excellent example of an atom with one electron in its outer ring. The positive charge on the nucleus of the copper atom is 29 as shown in Fig. 6. The first three rings of the atom are filled, they hold all the electrons they can. However, the 29th electron required to neutralize the charge on the nucleus is in the fourth ring by itself. This electron is not held very closely to the nucleus. As a result, it can move easily from one atom to another. This is the reason why copper wire is so widely used in electronic equipment and in electric power distribution.

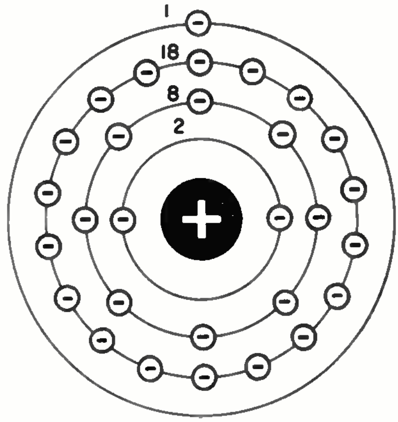


Fig. 6. The nucleus of the copper atom has a positive charge of 29. Around the nucleus there are normally 29 electrons arranged as shown above.

If we apply some external force to a copper atom, we can easily knock the outermost electron loose and it might move to an outer ring of a nearby atom. This atom will then have two electrons in the fourth ring. It then has one more electron than it needs to neutralize the charge on the nucleus. The tendency is for this atom to get rid of this extra charge as quickly as possible. Either the new electron that moved into the fourth ring will be forced out of this ring, or the original electron in the fourth ring will move out. In any case, whichever of the two electrons leaves this fourth ring will move over to a nearby atom, and it in turn will upset the balance of this atom and either move on itself, or force the electron in the outer ring of this atom out.

Now if you will remember when we rubbed the two glass balls and the two rubber balls with a silk cloth and placed like charges on them they repelled each other. Electrons have a negative electrical charge. All electrons have exactly the same charge. Since the charges on electrons are like charges they tend to repel each other.

In a piece of copper there will be millions of atoms. Each of these atoms will have a nucleus that has a positive charge of 29 on it and around it 29 electrons that neutralize this positive charge. The electrons are held in the atom by the positive charge on the nucleus which attracts them. At the same time the electrons are repelling other electrons in the atom and electrons in nearby atoms. There is more or less a condition where there is a balance between the nucleus holding or attracting its electrons and at the same time the electrons repelling or pushing away other electrons.

If we take a piece of copper wire and connect something to it that will try to pull electrons from one end of the wire and push them into the other end we will set up an instantaneous chain reaction along that wire. The instant an electron starts to move out of the fourth ring of one of the copper atoms the negative charge on

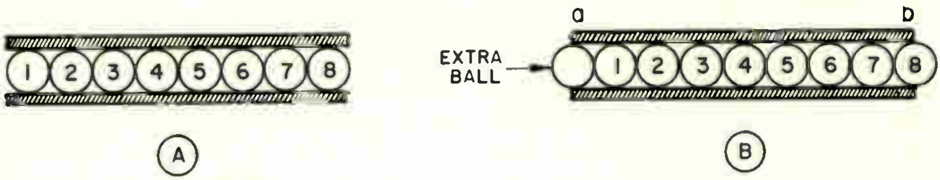


Fig. 7. The hollow tube shown at A is full of ping pong balls. When an extra ball is pushed into end a at B, the effect is to push all the balls at once and a ball starts to to fall out at end b.

that electron will push an electron out of the fourth ring of a nearby atom. At the same instant, it in turn will start pushing an electron out of the fourth ring of an atom adjacent to it. This will happen all the way along the wire so that at the instant an electron starts moving at one end of the copper another electron will start moving at the other end. The motion of the electrons will be the same all through the length of the wire.

You might get a better idea of what is happening if you took a hollow tube such as shown in Fig. 7A and filled it with ping pong balls so that the balls are all touching each other. The minute you start to force an extra ball into the one end, a, all the balls in the tube start to move and a ball starts to fall out of the end b as shown in Fig. 7B. The movement is instantaneous through the entire length of the tube.

The same situation exists when you start an electron moving at end a of the wire shown in Fig. 8. Although the electrons are not touching each other there is a force between them as shown, so that this force causes electrons all down the wire to start to move at the same instant. If you were to apply a greater force so that two electrons started moving at end, a, of the wire as shown in Fig. 9A, this force would cause two electrons to start moving all the way down the wire as shown. Similarly if you increase the force still further and start three electrons moving at end, a, as shown in Fig. 9B, then you have this chain reaction of three moving the entire length of the wire. The motion of electrons will be the same throughout the length of the wire.

The movement of electrons along the wire is called current flow. This is what an electric current is. We will go into this in detail in the next section of this lesson. However, be-

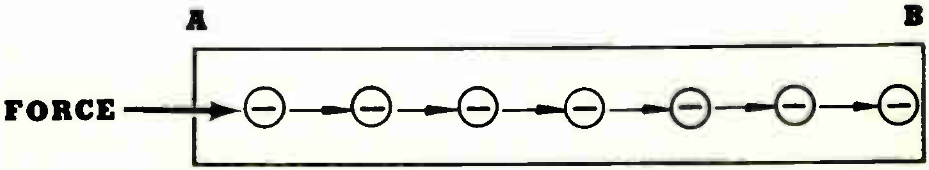


Fig. 8. When force is applied at end A of wire and an electron is pushed towards B, there is an instantaneous reaction all along the length of wire pushing a line of electrons towards B.

fore going to the next section let us explain the action of the charged rubber and glass balls. When the glass ball is rubbed with the silk cloth, the friction of rubbing the ball removes some of the electrons from the ball. Once the electrons have been removed, there are not enough electrons left to completely neutralize the charges on the nuclei of the atoms. Therefore the glass balls will have a positive charge on them.

When we rub the rubber balls, the rubber balls take electrons from the silk cloth so they will have more electrons than are needed to completely neutralize the positive charges on the nuclei of the atoms. Therefore the rubber balls will have a negative charge on them.

Since the two similarly charged glass balls repelled each other and likewise the two charged rubber balls repelled each other, we assumed that like charges repelled. Indeed, two positive charges will repel each other, and two negative charges will repel each other.

When we charge one glass ball and one rubber ball and suspend them near each other, they will be attracted because they have unlike charges; one is charged positive and the other is charged negative and unlike charges attract. When they move together and touch, some of the extra electrons on the rubber ball will leave the rubber ball and move over to the glass ball and partly make up for the shortage of electrons on the glass ball. When this happens, the balls may swing apart, but the charges will not be completely neutralized, so the balls will swing back together again and a few more electrons will move from the rubber ball over to the glass ball. This swinging back and forth will continue until enough electrons have moved from the rubber ball to the glass ball to reduce the force of attraction between the two balls until it is no longer strong enough to cause the balls to swing together.

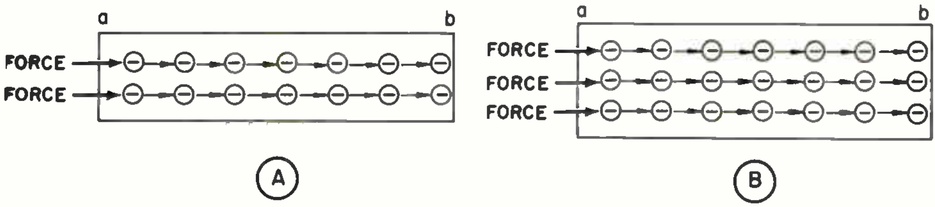


Fig. 9. Number of electrons set in motion along a wire depends on the force applied.

SUMMARY

We have covered a great deal of material in this section, and chances are you will not be able to remember all of it. We do not expect you to remember all of the details, but you should remember the important points such as:

1. Like electric charges repel, and unlike electric charges attract.
2. All material is made up of extremely small particles called atoms.
3. All the atoms of a given substance are identical.

4. Atoms are made up of a nucleus in the center, which has a positive charge, and a number of electrons, which have negative charges. Normally the atom will have enough electrons to exactly neutralize the charge on the nucleus.
5. The electrons arrange themselves in rings around the nucleus of the atom. There is a maximum number of electrons that can be in each ring.
6. In some atoms, such as the copper atom, an electron can be easily displaced.

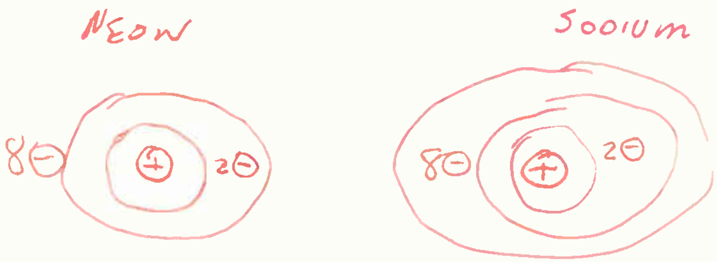
SELF-TEST QUESTIONS

Before going on with the next section of this lesson be sure to answer the following self-test questions. Write out the answer to each question carefully. After you have answered all of the questions, check your answers with those on pages 38 and 39. We do not expect you to give the same answer, but be sure that you understand the point brought out by the question before going on to the next section. Remember, do NOT send your answers to the Self-Test Questions to NRI for grading.

(a) State the law of charges.
(b) What is an atom?
(c) Which two parts of the atom are we interested in?
(d) Which part of the atom has a positive charge? Which part has a negative charge?
(e) If the copper atom which normally has 29 electrons loses one of its electrons, what kind of a charge does the atom have?
(f) Why do electrons in adjacent atoms repel each other?

If you can answer the preceding self-test questions you can be sure that you understand the important points covered in the preceding section of this lesson. However, before going on, here is a real tough question for you to think about. Don't spend more than five or ten minutes thinking about it, but try to answer it because this will stick in your mind and help you to remember some of these important points later.

(g) One atom has ten electrons and another atom has eleven electrons. Which of the two atoms would you expect would most readily give up an electron?



1⊖ The 11-electron atom

Current Flow

In the preceding section you have seen how an electron being knocked out of its atom forces additional electrons out of their atoms and sets them in motion. We have also pointed out that if we apply some force to push the electrons at one end of a wire and another force to pull the electrons from the other end we can start an instantaneous motion of electrons along the wire with the electrons moving from one end towards the other.

A device that is capable of doing this is a flashlight cell. A flashlight cell, by means of the chemical action that occurs inside it, will push electrons from one terminal and pull them into the other. If we connect a wire between the two terminals of a flashlight cell, electrons will immediately be set in motion through the wire and through the cell as shown in Fig. 10. Electrons will be moving from the one terminal of the flashlight cell through the wire and back to the other terminal of the cell, and through the cell back towards the terminal from which they started. This is called an electric circuit.

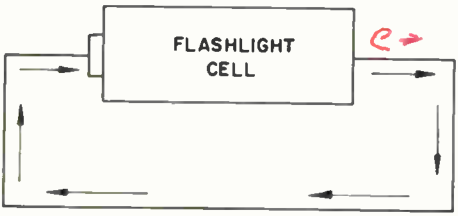


Fig. 10. A wire connected between the two terminals of a flashlight cell provides an electric circuit through which electrons can flow.

Electrons will continue to follow this circular path until the path or circuit is broken by disconnecting the wire from one terminal of the cell or until the chemical action of the cell is exhausted.

Notice that there must be a complete path for the electrons to travel. An electron leaving one terminal of the cell must be able to travel through a complete circuit, through the wire, through the cell and back to the terminal from which it left. If you simply connect a wire to one terminal of the cell, nothing will happen.

A flashlight cell has two terminals. One terminal is a small round terminal in the center of one end of the cell. This is the positive of the cell. The other end of the cell is the negative terminal of the cell. Now, can you tell from which terminal the electrons are going to leave the flashlight cell? Remember the experiment with the glass balls. When we charged the glass balls we placed a like charge on the two balls and they pushed each other apart. Now an electron has a negative or minus charge. Which terminal is going to push electrons away, and which terminal will attract them? The answer to these two questions is just what you might expect from the law of charges. We said that like charges repel, and therefore the negative terminal of the battery will push the electrons out of the battery. The positive terminal of the cell will attract the electrons, because it has the opposite charge on it.

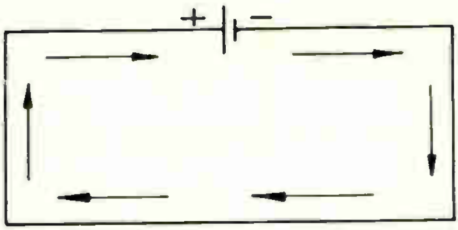


Fig. 11. The schematic diagram of a simple circuit.

In Fig. 11 we have repeated the circuit shown in Fig. 10; however, this time we have used the symbol that is used in electronics to identify a single cell such as a flashlight cell. Notice that we have used one short line and one longer line to represent the cell. The short line identifies the negative terminal of the flashlight cell and the long line the positive terminal. We have also marked the terminals with - and + signs so that there will not be any confusion. Remember the symbol used for the cell, you will see it many times in your electronics career.

Symbols like these are used on all electronic diagrams to indicate the various parts. There is a different symbol for each part. Using a simple symbol instead of trying to show the actual part makes the diagram much easier to understand. These diagrams using symbols are called "schematic" diagrams and the individual symbols, schematic symbols. We will teach you each symbol as you come to it so you will not have to learn a lot of them at once.

In the circuits shown in Fig. 10 and 11, the electrons leaving the negative terminal of the battery simply flow through the wire back to the positive terminal - they are not doing anything useful. Usually we will have some other device connected in the circuit so that the electrons flowing in the circuit will do something useful. For example, in a flashlight we have the flashlight bulb connected between the two terminals of the flashlight cell. The flashlight bulb is simply made of a piece of wire, usually some very hard wire such as a tungsten wire which is placed inside of the glass envelope from which all the air has been evacuated. When we connect the flashlight cell in the circuit such as shown in Fig. 12, the electrons will be set in motion instantaneously throughout the entire circuit as before. However, the movement of electrons through the tungsten wire produces a great deal of heat in the wire; in fact the wire gets so hot that it reaches "white" heat and gives off light. Here we have put the movement of electrons to work, and have used it to produce light. Circuits such as shown in Figs. 10 and 11 where the electrons simply travel from one terminal of the battery to the other are usually avoided. Sometimes they occur accidentally due to a parts failure and they are then referred to as short circuits. You will learn more about this in later lessons.

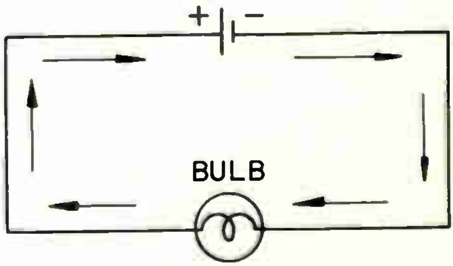


Fig. 12. A simple circuit showing a bulb connected across a flashlight cell.

THE AMPERE

The movement of electrons through an electric circuit is called an electric current. The strength of the current depends on the number of electrons in motion at any point in the circuit. As we pointed out previously, the number of electrons in motion will be the same at all points in the circuit. In the circuits shown in Figs. 11 and 12 the number of electrons leaving the negative terminal of the battery at a given time will be exactly equal to the number of electrons flowing through the flashlight bulb at the same time and also equal to the number of electrons reaching the positive terminal of the battery.

The number of electrons set in motion depends upon the force applied to the circuit, and also on the material used in the circuit. Some materials give up one or two electrons more readily than others, and as a result it is easier for the electrons to move through circuits made of these materials than in circuits made up of other materials that will not give up their electrons so easily.

We must have some way of knowing how much current there is flowing in a circuit. A movement of one or two electrons past a point in a circuit in a period of one second represents an extremely small current, so small in fact that it would be of no useful value. Before a current can be useful, there must be a tremendous number of electrons moving past each point in the circuit. It would be impractical to try to count the number of electrons, so instead, a unit of current called the "ampere" has been devised. The ampere represents a useful number of electrons flowing past a given point in the circuit per second. The actual number of electrons that will pass the point is unimportant. However, a standard ampere has been set up, and all current measurements are made in relationship to this standard ampere. If the number of electrons flowing in the circuit is twice the number represented by one ampere, then the current flowing in the circuit is two amperes. If it is ten times the standard ampere, the current flowing is ten amperes. The word ampere is used so often in electronics that we abbreviate it "amp". To make it plural, we simply add an "s", for example 10 amperes is written 10 amps.

THE VOLT

When we were discussing the ampere, we said that the amount of current that will flow in a circuit depends upon the force applied to the circuit. We should have some means of measuring this force. The force is called the "electromotive force", or "voltage", and it is measured in volts. Often you will see electromotive force abbreviated "emf".

You do not have to be concerned about exactly how much force there is in one volt; the important thing is to know that the number of volts indicates the amount of force applied to the circuit, and the higher the voltage, the more force is being applied to the circuit. In other words, two volts represents twice as much force as one volt. Ten volts represents ten times as much force as one volt.

By way of interest you might like to know that the voltage of a conventional flashlight cell is approximately one and a half or 1.5 volts. The storage batteries used in modern automobiles are made up of six cells connected so that the voltage of the six cells adds. Each cell has a voltage of about 2 volts so that the total battery voltage is 12 volts. Older automobiles may have 3 cell storage batteries - the voltage of these cells is 6 volts. Electric light bulbs and most appliances in homes are designed to operate on a voltage of about 120 volts.

THE OHM

The amount of current that will flow in a circuit depends on one other thing besides the force applied to the circuit. This is how readily the material will give up electrons and let them move in the circuit. Some materials will give up electrons quite readily and let them move through the circuit with little or no opposition. However, other materials will not give up electrons so readily, and may offer considerable opposition to the flow of current. This opposition to current flow is called "resistance". The resistance of a material depends upon how readily it will allow electrons to move through it. Resistance is measured in "ohms". Again, you need not know the exact definition of a standard ohm, the important thing to know is that resistance is the opposition to current flow in a circuit and that it is measured in ohms.

V=IR OHM'S LAW E=IR

We have said that the current that flows in a circuit depends upon the force or voltage applied to the circuit and on the opposition or resistance in the circuit. This means that the current depends both on the voltage and on the resistance.

If in an electrical circuit a voltage of 1 volt is applied to a circuit having a resistance of 1 ohm, a current of 1 ampere will flow in the circuit. If we double the voltage so that the voltage is 2 volts and the resistance is still 1 ohm, the current that will flow in the circuit will be 2 amps. On the other hand, if the voltage is 1 volt, and we double the resistance to 2 ohms, the current that will flow in the circuit will be only 1/2 amp.

This relationship between current, voltage and resistance is known as "Ohm's Law". We will use Ohm's Law many times in future lessons. For the present, all you need to remember is that the current depends upon the voltage and the resistance. If you double the voltage and keep the resistance constant, the current will double. If you cut the voltage in half and keep the resistance constant the current will be cut in half. On the other hand, if you keep the voltage constant but double the resistance the current will be cut in half but if you keep the voltage constant and cut the resistance in half the current will double. Increasing the voltage increases the current, reducing the voltage reduces the current. Increasing the resistance reduces the current and reducing the resistance increases the current.

SUMMARY

Here are the important points you should remember from this section of the lesson:

1. If a flashlight cell is connected to a wire made up of a material from which some of the electrons can be displaced, electrons will move from the negative terminal of the cell through the wire to the positive terminal, and through the cell back to the negative terminal making a complete circuit. The

cell provides the force that sets the electrons in motion.

2. A movement of electrons through the circuit is called a "current". In a simple circuit such as shown in Figs. 10 and 11, the movement of electrons is the same at all points in the circuit. Also, remember that once the wire is connected to the two-battery terminals and the electron movement starts, it starts the same instant in all parts of the circuit.
3. The unit used to measure the strength of an electric current is the ampere.
4. The unit of force that sets the electrons in motion is the volt.
5. The unit of resistance is the ohm.
6. The relationship between current, voltage and resistance is known as Ohm's Law.

SELF-TEST QUESTIONS

(h) Which terminal of a flashlight cell has a surplus of electrons? Which terminal has a shortage of electrons?

(i) When an electric circuit is completed do the electrons start in motion at the negative terminal of the battery first? At the positive terminal of the battery? Or do they start in motion instantaneously throughout the entire circuit?

(j) What is the unit in which we measure electric currents?

(k) What is the unit used to measure the electromotive force applied in an electric circuit?

(l) What unit is used to measure the opposition to current flow in an electric circuit?

(m) According to Ohm's Law, what effect on the electric current flowing in the circuit will increasing the voltage have? INCREASED

(n) If we reduce the resistance in an electric circuit, will the current in the circuit increase or will it decrease?

Magnetism

Magnetism is as important in electronics as electricity. Without magnetism there would be no electronics industry at all, for magnetism and electricity work together to make our modern electronic devices possible.

PERMANENT MAGNETS

A magnet will pick up or attract small pieces of steel or iron. Minerals that have this property are found buried in the ground in some parts of the world. These minerals are called natural magnets. A piece of iron or steel can be made into a magnet by stroking it in one direction with a magnet. When we make a magnet in this way, we say that the metal is "magnetized". These magnets are called "permanent" magnets, because they will retain their magnetism almost indefinitely.

Originally, permanent magnets were made of iron or steel, but modern permanent magnets are usually made of an alloy called "Alnico". Alnico is a mixture of aluminum, nickel and cobalt. Very strong lightweight magnets, which retain their magnetism much better than magnets made of iron or steel, can be made from Alnico.

When a magnetized steel needle is suspended at its balance point by a light thread, as shown in Fig. 13, the needle will always line up in a direction corresponding closely to north and south. This phenomenon led to the first practical use for magnets, in compasses used by early sea voyagers and travellers. Compasses are made simply of a magnetized piece of steel that is mounted on a delicately pivotted bearing that will turn as easily as the needle suspended by the thread.

POLES OF A MAGNET

The ends of a permanent magnet are called "poles". This name was given to the ends of a magnet because the ends point toward the poles of the earth when the magnet is free to pivot on an axis. The pole that points toward the north pole of the earth was originally called the "north-seeking" pole. However, for simplicity the name has been shortened to the "north pole". The magnetic pole that points to the south pole of the earth is called the "south pole".

If two magnets are brought near each other, the north pole of one

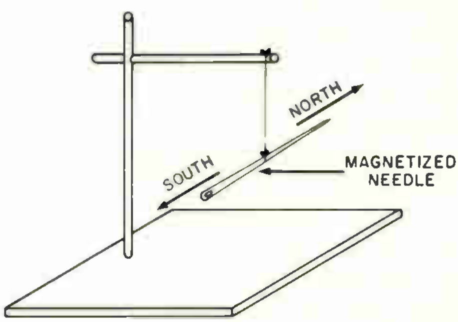


Fig. 13. If a magnetized needle is suspended at its balance point by a thin thread, the needle will line up in a north and south direction like a compass, because it lines up with the magnetic field of the earth.

magnet will repel the north pole of the other. Similarly, the south pole of one magnet will repel the south pole of another magnet. However, the north pole of one magnet will attract the south pole of the other magnet. The reason why a compass always points in a north-south direction is that the earth itself is a magnet, and one pole of this large magnet is near the north geographic pole, and the other pole is near the south geographic pole. The north pole of the earth attracts one of the poles of the magnet and repels the other. The south pole of the earth attracts the pole that is repelled by the north pole, and repels the pole that is attracted by the north pole. Therefore, the magnet will point in a direction so that one pole points toward the north magnetic pole and the other pole points toward the south magnetic pole.

Notice the similarity between the attraction and repulsion of magnetic poles and the attraction and repulsion of electric charges. You already know that "like charges repel and unlike charges attract". In magnets, "like poles repel and unlike poles attract". This is a fundamental law of magnetism; you should remember it.

MAGNETIC LINES OF FORCE

There are lines of force surrounding a magnet. You can trace out the lines of force around a magnet by using a small compass. If you bring the compass near the north pole of the magnet, the south pole of the compass will be attracted to the north pole of the magnet. The compass needle will line up with the magnetic lines of force. If you move

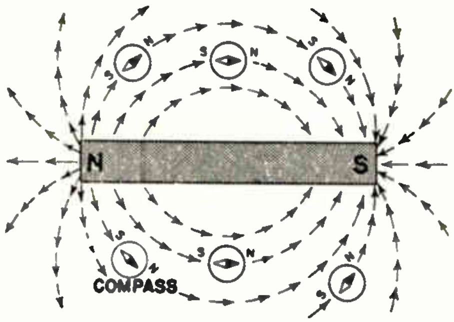


Fig. 14. A small compass can be used to trace magnetic lines of force near a permanent magnet.

the compass as shown in Fig. 14, you will be able to trace out the lines of force. The lines of force are shown coming from the north pole of the magnet and going to the south pole. We do not know for sure whether or not this is true, but there is considerable evidence that the lines of force actually do go from the north pole to the south pole and therefore we will base our explanations on this assumption.

Another experiment that can be performed to show the lines of force around a magnet is to place a thin sheet of cardboard over a magnet and then sprinkle iron filings evenly over the sheet of cardboard. Tap the cardboard gently, and the iron filings will arrange themselves in definite lines, producing the pattern shown at the top in Fig. 15.

If two magnets are arranged so that their north poles are placed close together and then iron filings are sprinkled on a cardboard placed over the magnet, the iron filings will arrange themselves as shown in the center of Fig. 15. Notice that you can see the lines of force from the north poles of the two magnets actually repelling each other.

You would get a third pattern by

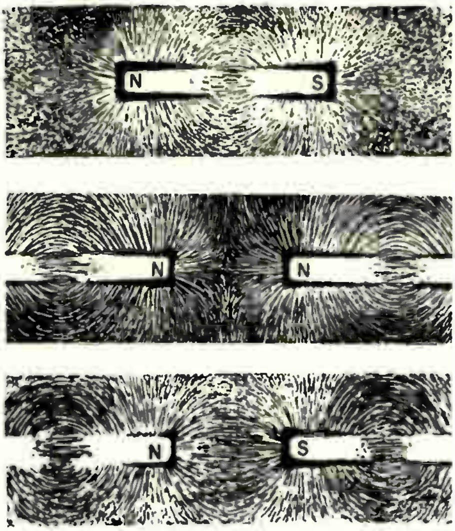


Fig. 15. If iron filings are placed on a sheet of cardboard over permanent magnets, they will trace out lines of force as shown here.

placing the north pole of one magnet toward the south pole of another and sprinkling iron filings on a cardboard. The pattern you would get in this case would be like the one shown at the bottom in Fig. 15. Here you can see the attraction between the north pole and the south pole of the two magnets.

The lines of force coming from a magnet are called "magnetic lines of force". They are similar to the lines of force surrounding electrically charged objects. The lines of force around an electrically charged object are called "electric lines of force".

ELECTRO MAGNETS

An electric current flowing through a wire produces a magnetic field in the space around the wire. This is called "electro magnetism". The circular lines of force around the wire can actually be traced out with a compass. There will be many of these magnetic rings surrounding the entire length of the wire. The magnetic lines of force close to the wire will be much stronger and more easily detected than those at some distance from the wire as shown in Fig. 16. However, even with a weak current flowing through a wire, magnetic lines of force can be detected some distance from the wire, with sensitive equipment.

Even though the magnetic lines of force around a current-carrying wire can be detected, the magnetic field will be weak unless the current flowing through the wire is very strong. However, by winding the wire in the form of a coil, a strong magnet can be made. When the wire is bent into a loop, the circular magnetic rings pass through the center of the coil in the same direction and reinforce each other as shown in Fig. 17.

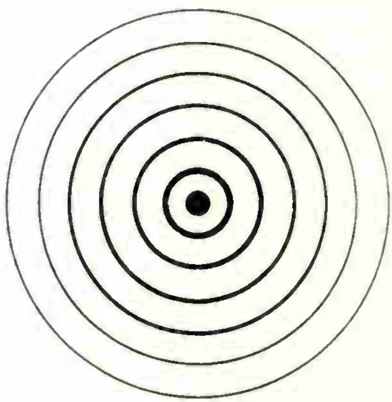


Fig. 16. When electrons flow through a wire, magnetic lines of force surround the wire. This is a cross-sectional view.

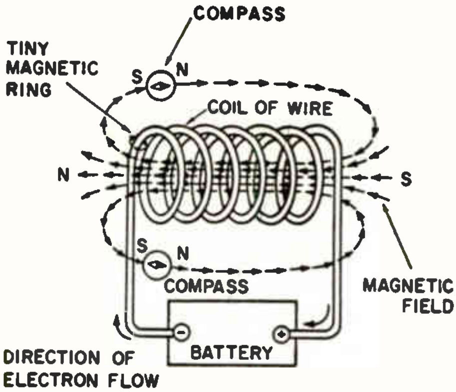


Fig. 17. The more turns of wire we have in a coil, the stronger the magnetic field will be when electrons flow through the coil.

This type of magnet is called an "electromagnet". The magnetic effect exists only as long as the current is flowing through the wire. Once the current is stopped by opening the circuit, the magnetic effect will disappear. Many parts used in electronic equipment depend upon the basic principles of electromagnetism to operate.

The electromagnet shown in Fig. 17 can be made much stronger by inserting an iron bar or a bar of some magnetic material inside the coil. The iron bar is called a "core". The actual increase in the strength of the magnet will depend upon the type of core material used.

You probably wonder why inserting a core inside the coil makes the magnet stronger. The answer to this question is that the iron core is made up of millions of tiny particles of iron. Each of these particles is itself a magnet having a north pole and a south pole. Ordinarily, these tiny magnets are not arranged in any definite pattern. One might point in one direction and another in a second direction, and a third in still another direction as shown in Fig. 18A. As a result of the random arrangement of these small magnets, the magnetic field of one magnet is cancelled by the magnetic field of another. However, when the iron core is placed in the magnetic field inside the current-carrying coil, the magnetic field produced by the coil causes the particles to line up and all point in the same direction as shown in Fig. 18B. When this happens, the entire bar becomes one strong magnet. However, most of the tiny particles are kept lined up only by the magnetic field produced by the current flowing in the coil. Once this field is removed by opening the circuit so that the current can no longer flow through the coil, most of the tiny particles will return to their random arrangement so that they will no longer be pointing in one direction, and most of the magnetic field will disappear.

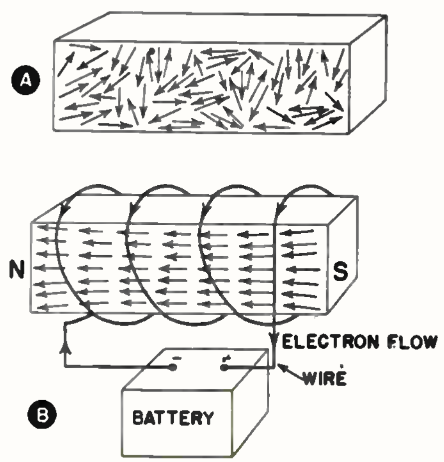


Fig. 18. In an unmagnetized bar of iron at A, the tiny magnets in the iron do not line up; if the bar of iron is magnetized, they will line up as at B.

INDUCED CURRENTS

We have seen that there is a magnetic field around a current-carrying wire, and that if a current flows through a coil, an electromagnet will be produced. Now, is the opposite true? If a coil is placed inside a magnetic field will a current flow through the coil? Let us look in the experiment illustrated in Fig. 19. Here we have a coil wound on a hollow form. The ends of the coil are connected to a small flashlight bulb. We have used a combination of a pictorial drawing for the coil and a schematic symbol for the bulb.

If a magnet is moved quickly inside the hollow form, the bulb will light while the magnet is being moved into the coil. Once the magnet is completely inside the coil and no longer moving, the bulb will no longer light. When the magnet is moved quickly out of the coil, the bulb will light again. If the magnet is moved quickly in and out of the coil, the bulb will light and remain lighted as long as the magnet is in motion. As long as the magnet is moving inside the coil, a current will flow in the coil and through the flashlight bulb.

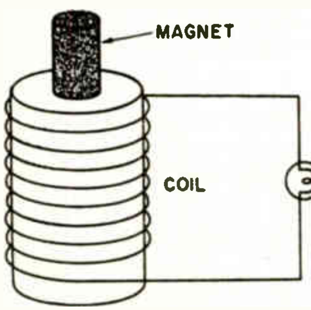


Fig. 19. If a magnet is moved in and out of a coil connected to a flashlight bulb, the magnetic lines of force cutting the turns of wire on the coil will induce a voltage in the coil, and the voltage will cause a current flow through the flashlight bulb.

This current flows because a voltage is induced in the coil. The magnetic lines of force moving through the turns of wire on the coil are said to cut the turns of wire on the coil. A small voltage is induced in each turn on the coil as long as the number of magnetic lines of force cutting the turn is changing. The voltages induced in the various turns of wire on the coil add together. This total voltage produces a current flow through the coil and through the flashlight bulb. We call the voltage produced an "induced" voltage and the current an "induced" current.

A second demonstration of an induced voltage is shown in Fig. 20. Two coils are wound on the same form and placed near each other. It is customary on schematic diagrams to use letters to designate the various parts. L is usually used for coils. We have marked the coils L1 and L2 to make them easy to refer to. One coil, which we have marked L1, is connected to a flashlight cell through a switch, and the other coil marked L2 is connected to a flashlight bulb. Inserted through the form on which the coils are wound is an iron core to increase the strength of the magnetic field. We have shown the schematic symbol for the switch, and labelled it SW. This makes three schematic symbols you should know now, the flashlight cell, the light bulb and the switch.

When the switch is closed and current starts to flow in L1, there will be a voltage induced in L2 and a glow will be seen in the flashlight bulb. However, this glow will last for only

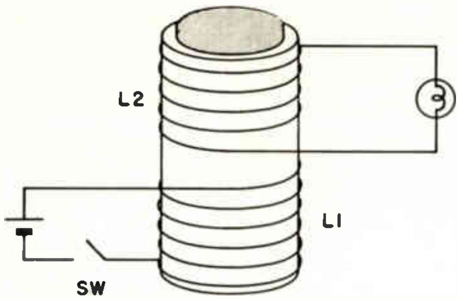


Fig. 20. When the circuit to L1 is opened or closed, a voltage will be induced in L2.

an instant after the switch is closed. When the switch is open and the current flow in L1 is interrupted, the bulb again will glow, indicating that this too induced a voltage in the second winding.

The explanation of this action is the same as for the voltage induced by the magnet moving in and out of the coil in Fig. 19. When the circuit is completed by closing the switch, current starts to flow in the coil. But the magnetic field that accompanies the current flow does not build up instantly; it takes just a short time for this field to build up. When the field is building up, the magnetic lines cutting the turns of L2 are changing just as they cut the coil when the magnet was moved into the coil in Fig. 19. While the magnetic field in L1 is building up, the number of magnetic lines of force is increasing. This change in the number of magnetic lines of force cutting L2 induces a voltage in L2 which causes a current to flow through it and the flashlight bulb. When the circuit is open, the magnetic field must disappear. The number of magnetic lines of force again must change and while this is happening, the change in the number of magnetic lines of force cutting L2 again induces a voltage in L2 which causes the current to flow.

It is important for you to realize that whenever a magnetic field around a coil changes, there will be a voltage induced in the coil. Remember this, it is important; you will be dealing with induced voltages as long as you are in the field of electronics.

SUMMARY

There is a great deal of similarity between magnetism and electricity. Indeed, the basic law of magnetic forces - "like forces repel, and unlike forces attract", is the same as the basic law of electricity, "like charges repel and unlike charges attract". As you go on through your course you will see that electricity and magnetism work together to make the electronic devices that we use today possible.

You are not expected to remember all the details described in this section on magnetism. The information is presented so you will have a complete picture and be able to understand magnetism completely. The important points that you must remember from this section are as follows:

1. A magnet has a north pole and a south pole.
2. Magnetic lines of force travel from the north pole of the magnet around through space to the south pole of the magnet.
3. Like magnetic poles repel; unlike magnetic poles attract.
4. There is a magnetic field around a current-carrying wire.
5. An electromagnet can be made by passing a current through a coil.

6. Inserting an iron core into an electromagnet will result in a stronger magnetic field.
7. If the magnetic lines of force cutting a turn of a coil change, there will be a voltage induced in that turn of the coil.
8. If the magnetic lines of force cutting all the turns of a coil change there will be a voltage induced in each turn of the coil and these voltages will add together. If the coil is connected into a complete circuit, current will flow in the circuit.

SELF-TEST QUESTIONS

(o) State the basic law of magnetism.

(p) If a pole of a magnet attracts the south pole of a compass, is the pole a north or south pole?

(q) In which direction do the magnetic lines of force coming from the poles of a magnet travel?

(r) Where are the magnetic lines of force around a current-carrying wire strongest - close to the wire or at some distance from the wire?

(s) What effect will placing an iron core inside of an electromagnet have on the strength of the magnetic field?

(t) If the ends of a coil are connected to a flashlight bulb, and a very strong permanent magnet is placed inside of the coil, will the flashlight bulb light? How long will the flashlight bulb remain lit?

(u) In the circuit shown in Fig. 20, when the switch is first closed, the flashlight bulb will glow. Why doesn't the flashlight bulb continue to glow as long as the switch is closed?

Electronic Components

The parts used in electronics equipment are often simply called parts, but they are sometimes called components. The two words mean the same thing; we'll use both words so you will get familiar with them.

In the rest of this lesson, you will study a few of the parts found in electronic equipment. We are simply going to introduce these parts; you will study them in detail in later lessons.

You have already seen the schematic symbol for a single-cell battery such as the flashlight cell, for a light bulb and for a switch. You will learn the symbols used for several other parts. It is important that you learn these symbols as you go along. The symbols are used to draw schematic diagrams. Schematic diagrams tell you how the various parts are connected together. You must learn how to read this type of diagram. Manufacturers supply schematic diagrams of electronic equipment, they do not supply picture diagrams. As a matter of fact, picture or pictorial diagrams would be far more complicated and far more difficult to read than schematic diagrams. Once you learn how to read schematic diagrams you will find that they tell you far more about how a circuit is connected than a pictorial diagram could possibly do.

It is quite a job to learn all the schematic symbols at once, but if you take them one at a time, as you come to them in your lessons, and also learn how to read the simple schematic diagrams that will be shown in the early lessons, you will soon find that you know the symbols and that schematic diagrams are really quite easy to follow. If you learn how to read the simple schematic diagrams in the beginning, when you get along further in your course you will find that the large complex diagrams, that you will have to deal with later, are very easy to follow.

CONDUCTORS AND INSULATORS

When we connect parts together in an electronic circuit there will be certain paths through which we want electrons to flow, and other places where we want to avoid a current flow. We use conductors to provide the paths for current flow and insulators where we want to prevent current flow.

There are a number of materials from which one or two electrons in the outer ring of electrons can be displaced. Copper, silver, aluminum, iron and most metals are examples of this type of material. Since electrons can be displaced from these materials it is easy to set them in motion in a wire made of this type of material, and cause a current to flow through the wire. These materials are called conductors. They are so called because they will conduct or transmit an electric current; in other words a current flow can be set up in them. The lines used to connect parts together on schematic diagrams represent the wire conductors used to provide paths for current flow.

There is no such thing as a perfect conductor. All conductors offer some resistance, or opposition, to an electric current. Silver is the best known conductor, but it is used only in special applications because it is too expensive. Copper, which is almost as good a conductor as silver, is used in most wire because it is less expensive than silver. Copper wire is used almost exclusively in connecting electronic components together. It is also used in coils. However, in some special applications where it is essential to keep the resistance as low as possible you will find that silver wire or silver plated copper wire is used. You will learn about these special applications later.

There are other materials which will not readily give up any electrons. A material having its outer ring full of electrons has no place to permit any additional electrons to move into the atom nor will it willingly permit any electrons to move from any of its rings. This type of material is called an insulator. It normally will not pass or conduct an electric current.

There is no such thing as a perfect insulator. Even in materials having all the electron rings filled, an electron will occasionally escape, particularly if enough force is applied to the material. However, the number of electrons that will escape is usually so small, that for all practical purposes we can say that these materials will not conduct current. When an extremely high force is applied to this material, electrons may be forced out and the material will break down, and no longer will be usable as an insulator.

Most copper wire that you will use in electronic equipment to connect the various parts together will be covered with a rubber coating or a plastic coating. The coating is an insulator. Its purpose is to keep the current flowing through the wire to the part it is supposed to reach and prevent its travelling through another circuit. If we did not use an insulator over the wire, and two wires happened to accidentally touch, we might have a short circuit where current would simply flow out one wire and back to the battery and perform no useful service. With an insulating material around the wires, the insulator will prevent this from happening.

BATTERIES

You already know that a flashlight cell is a device that can force electrons through a circuit. You know that the cell has two terminals, a positive terminal and a negative terminal. You know that when the flashlight cell is connected into a circuit, electrons will leave the negative terminal of the cell, flow through the circuit and back to the positive terminal of the cell. You also know that the voltage of a flashlight cell is about 1.5 volts.

Sometimes we need more voltage than can be obtained from a single flashlight cell. For example, you have surely seen a flashlight in which two cells are used. In such a flashlight the cells are arranged as shown in Fig. 21A. The positive terminal of one cell connects to the center terminal on a flashlight bulb. The positive terminal of the second cell is connected to the negative terminal of the first cell. The negative terminal of the second cell connects to a switch and through the switch to the threaded part of the bulb. When the switch is closed current flows

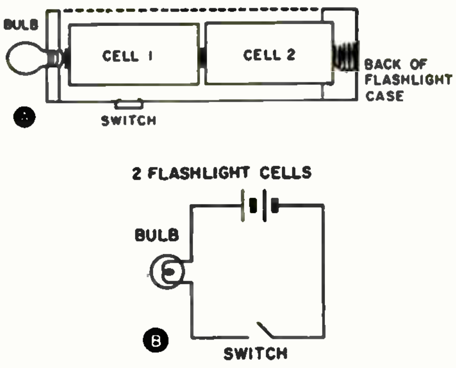


Fig. 21. Two flashlight cells used in series to power a 3-volt flashlight bulb are shown at A. The schematic diagram of this circuit is shown at B.

from the negative terminal of cell 2 through the switch and then through the bulb to the positive terminal of cell 1, through cell 1 to the negative terminal of cell 1 and across to the positive terminal of cell 2 and then through cell 2. Schematically the circuit is shown in Fig. 21B.

In a circuit of this type we say that the two cells are connected in series. Each cell provides a voltage of 1.5 volts so that the total voltage applied to the flashlight bulb is 1.5 + 1.5 or 3 volts.

You can obtain devices in which two cells similar to a flashlight cell are put in a single container to provide a total output voltage of 3 volts. When two cells are put together like this we call it a battery. In other words, a battery is simply a device in which there are several cells. Often we call a flashlight cell a flashlight battery; technically this is not quite correct, but it has come into such wide usage that everybody knows what is meant and as a result the expression is used.

In some applications in electronics you might need a voltage of 4.5 volts. To get this voltage all you need to do is connect three 1.5 volt flashlight cells in series and the voltages will add to give you a voltage of 4.5 volts.

Many of the small portable transistor radios in use today use a small 9-volt battery. Batteries of this type are simply made up of six small cells similar to the flashlight cell. Six times 1.5 gives you a voltage of 9 volts.

In the early days of radio 22.5 volts and 45-volt batteries were widely used. A 22.5 volt battery had fifteen 1.5 volt cells connected in series to give a voltage of 22.5 volts and a 45-volt battery simply had thirty 1.5-volt cells connected in series to give a voltage of 45 volts.

A 3-volt battery is generally indicated schematically by the symbol for two cells arranged such as shown in Figs. 21B and 22A. A 4.5-volt battery usually used three cell symbols as shown in Fig. 22B. However, if we wanted to show a 22.5 or 45-volt battery it would be too tedious to draw the symbol for the required number of cells so the symbol usually shows 5 or 6 cells connected in series such as shown in Fig. 22C and then the voltage is written either above or below the cell as indicated in the figure.

When cells are connected so that

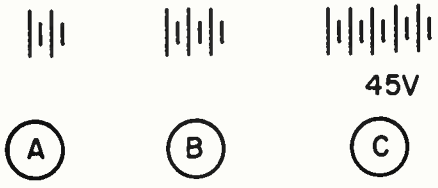


Fig. 22. Schematic symbol for two cells in series is shown at A. Symbol for three cells in series is shown at B. C is symbol for 45V battery.

the negative terminal of one is connected to the positive terminal of another to produce a battery, the total battery voltage is equal to the voltage of the individual cells times the number of cells. However, sometimes a number of cells are connected so that the positive terminals of all the cells are connected together and the negative terminals of the cells are connected together. When cells are connected in this way, we also refer to the device as the battery. However, a battery of this type has an output voltage equal to the voltage of only one cell. We say that the cells are connected in parallel.

You might wonder why we would want to connect cells in parallel. The answer is that in some applications we may need more current than can be supplied by a single cell. In this case by connecting a number of cells in parallel, each cell can supply part of the required current and the total number of cells is connected together to form a battery that is capable of supplying the current required.

This might immediately bring up a question - why not simply make a bigger cell that is capable of supplying the current needed. This can be done, but often manufacturers are making certain size cells in very large quantities and therefore they can make them at a low cost. To make a single cell that could supply two or three times the current capacity of a single cell might cost as much as ten or fifteen times what it would cost to make the smaller cell that they were making in very large quantities. Therefore it is more economical to take three or four of these smaller cells and connect them in parallel for applications where high currents are required than it would be simply to make one special cell that could supply the currents and have to make them only in limited quantities.

When a number of cells are connected in parallel and the output voltage is only 1.5 volts we usually use the same schematic symbol as we use for a flashlight battery. This symbol indicates the voltage, it does not indicate that the cell is capable of a higher current than a flashlight cell. If we want to make it clear that we have several cells connected in parallel, we can use the symbol shown in Fig. 23. This symbol shows four cells connected in parallel.

Other Types of Batteries.

The flashlight cell is only one type of cell. There are many other types of cells. For example, there is the lead cell which is used in storage batteries found in automobiles. Six lead cells are arranged in series in the average automobile battery found in late model cars. In older cars three cells were connected in series. The lead cell has a voltage of about 2 volts so that modern cars have a 12-volt battery in it whereas older cars have 6-volt batteries.

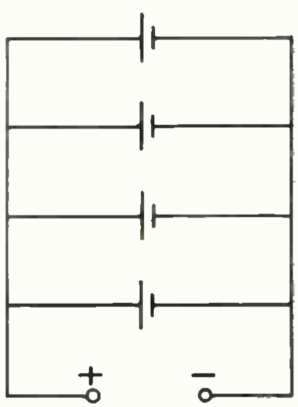


Fig. 23. Four cells connected in parallel.

Other types of cells used today are the mercury cell and the manganese cell. You will study these cells and batteries made up of these cells in later lessons.

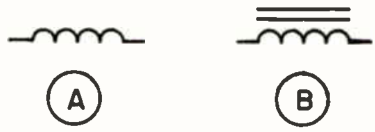


Fig. 24. Schematic symbols for coils. The lines above the coil at B indicate an iron core.

COILS AND TRANSFORMERS

You have already been briefly introduced to coils and you know that if the number of magnetic lines of force cutting a coil changes, a voltage will be induced in the coil. In electronics you will run into all kinds of coils. In the tuners of television receivers designed to receive the ultra high frequency channels, you will find coils that have only one or two turns. In other applications you will find coils having many turns. The schematic symbol used to represent a coil is shown in Fig. 24A. This should not be too hard to remember because the symbol itself looks something like a coil.

We mentioned previously that sometimes an iron core is placed inside the coil, and that placing the iron core inside of the coil will greatly increase the magnetic field produced by the coil. Often in electronics there will be iron cores used inside of a coil. When a coil has an iron core, a schematic symbol like that shown in Fig. 24B is usually used. The lines placed beside the coil symbol indicate that the coil has an iron core.

Probably no device has done more for the electronics industry than the transformer. The transformer has made it possible for power companies to supply homes and industry with electric power economically. Without economical power there could be no electronics industry. There is hardly a piece of electronic equipment made that does not use one or more transformers.

In spite of the importance of the transformer, it is basically a simple device. A transformer in its simplest form is nothing more than two coils mounted close together. The two coils we discussed in Fig. 20 actually can be called a transformer. Two typical transformers and the schematic symbols for them are shown in Fig. 25. The transformer shown in Fig. 25A consists of two coils wound on a cardboard frame. This type of coil is called an air-core transformer. The one shown in Fig. 25B is made of two coils wound on iron core. This type is called an iron-core transformer.

Air-core transformers such as shown in Fig. 25A are used in radio frequency applications. By radio frequency we mean radio signals. Iron-core transformers such as shown in Fig. 25B are used in power

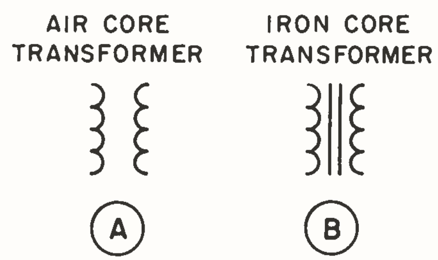


Fig. 25. Schematic symbols for transformers. The transformer at A is an air-core transformer, the one at B, an iron-core transformer.

applications. You will find out what the difference is between radio frequency signals and power frequencies in your next lesson.

We do not expect you to know all there is to know about coils and transformers at this time, the only thing we want you to remember is the schematic symbols used for coils and the schematic symbols used for transformers. You will go into great detail on these important parts in a later lesson.

CAPACITORS

An important electronic part is the capacitor. Basically a capacitor is simply two metal plates that are placed close together. The plates do not touch, they may be separated simply by an air space or some other material may be placed between the two plates of the capacitor.

If a battery is connected to the two plates of the capacitor, the negative charge on the negative terminal of the battery will try to force additional electrons into the one plate of the capacitor. These electrons will repel electrons from the other plate of the capacitor and they in turn will flow towards the positive plate of the battery. As a result, we will build up a charge on the two plates of the capacitor as shown in Fig. 26A. Notice that one plate of the capacitor has a negative charge and the other plate has a positive charge. The schematic symbols used for a capacitor are shown in Fig. 26B. You will find both types of symbols used. As you might expect the two lines represent the two plates of the capacitor. Be sure that you remember the schematic symbols. Capacitors are among the most important parts used in electronics. There are many different sizes and different types; capacitors are so important that we will devote an entire lesson to them and to their uses later.

RESISTORS

Earlier in this lesson we mentioned that conductors were used to carry the electric current from one

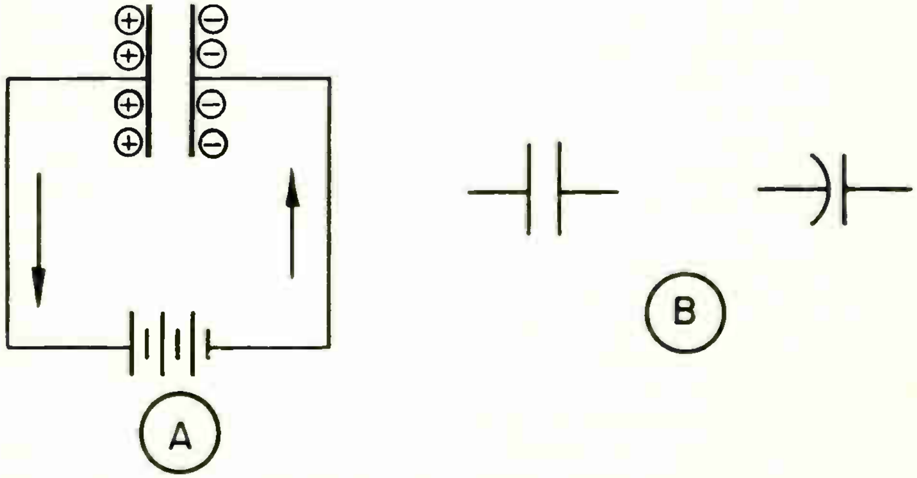


Fig. 26. The circuit at A shows how a capacitor can be charged. The schematic symbols are shown at B.

part of a circuit to another. We pointed out that the materials used in conductors were selected because they had electrons readily available and offered little or no opposition to the flow of electric current through them. In some applications we want to offer opposition to the flow of electric current. In these cases we use a device called a resistor. A resistor may be made of a carbon-type composition or it can be made of a wire that does not have as good conduction capabilities as copper has. In electronics you will run into resistors having a resistance of only a few ohms up to resistors that may have a resistance of well over 1,000,000 ohms.

Fig. 27. The schematic symbol for a resistor.

The schematic symbol for a resistor is shown in Fig. 27. The actual resistance of the resistor is usually written beside the schematic symbol as in Fig. 27. Here we have a resistor that has a resistance of 220 ohms and we have indicated this value above the resistor.

In any piece of electronic equipment you will probably find more resistors, capacitors and coils than any other parts. As a result, their schematic symbols will appear most frequently on schematic diagrams.

Be sure you remember the symbols used for each of these three important parts. Resistors are so important and so widely used that you will go into detailed study of them in a later lesson and you will be dealing with them throughout your entire electronics career.

VACUUM TUBES

Vacuum tubes are so widely used today that almost everyone has seen one. Since tubes are so widely used, it is important for you to learn something about their operation as soon as possible.

The Diode Tube.

The simplest vacuum tube is the diode tube. In the diode tube a filament that can be heated to a red heat by passing a current through it is placed inside of a glass envelope. Around the filament is a metal cylinder called the plate. Leads are brought out of the glass envelope for the two filament leads and for the plate lead. All the air is evacuated from the inside of the glass envelope before the envelope is sealed. The schematic symbol for a diode tube is shown in Fig. 28.

We mentioned earlier that in an atom, the electrons were rotating about the nucleus of the atom. In a diode tube the electrons are rotating about the nucleus of the atom in the material used for the filament. When a battery is connected between the filament terminals as shown in Fig. 29A, the filament is heated to a red heat. This causes the motion of the electrons to speed up and many of

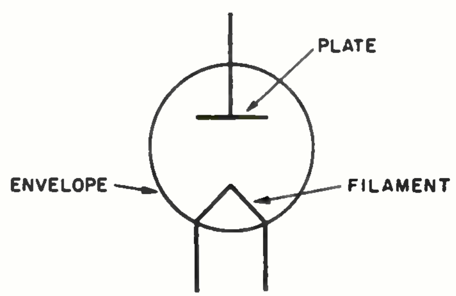


Fig. 28. The schematic symbol of a two-element (diode) tube.

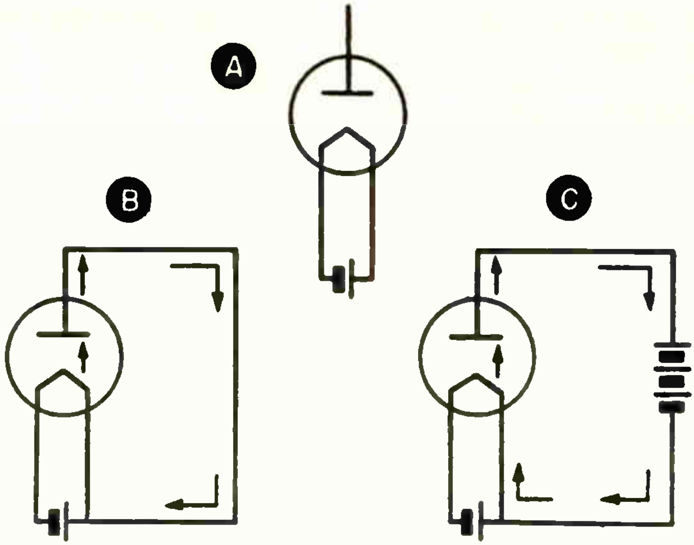


Fig. 29. The filament of a diode is heated as shown at A. When a diode is connected as shown at B, a small current will flow in the direction indicated by the arrows. When a battery is added as in C, a much stronger current will flow.

the electrons to break loose from the atom and fly off into space around the filament of the tube. If we connect the lead from the plate of the tube back to the filament as shown in Fig. 29B, some of the electrons that fly off the filament will travel through the space from the filament of the tube over to the plate and then flow from the plate through the external circuit back to the filament of the tube.

Since electrons have a negative charge they are attracted by a positive charge. Therefore if we connect a battery between the plate and the filament as shown in Fig. 29C, the electrons that fly off the filament of the tube will be attracted by the positive potential on the plate of the tube. As a result, many more of the electrons will travel from the filament over to the plate of the tube to the positive terminal of the battery. Electrons will travel through the battery to the negative terminal and from the negative terminal back to the filament of the tube.

Of course, as with any complete circuit, the current flow around the circuit is instantaneous. The instant the battery is connected to the tube current starts flowing around the circuit, and the amount of current flowing through the circuit is the same at all points in the circuit at all times.

Even though the diode tube is the simplest and the first tube invented, it is still in use today. The high-voltage rectifier used in television receivers today is nothing other than an improved version of this simple diode tube. Be sure that you remember the schematic symbol for the diode tube that is shown in Fig. 28 - this is an important symbol and you must remember it.

The Triode Tube.

While the diode tube is important,

and its discovery was a great milestone in the early days of electronics, it was not until the three-element tube called the triode tube was invented that the electronics industry as we know it today really got started. The schematic symbol of a triode tube is shown in Fig. 30. Notice that the symbol is the same as the symbol for the diode tube except that a third element has been added between the filament and the plate. This third element is called a grid.

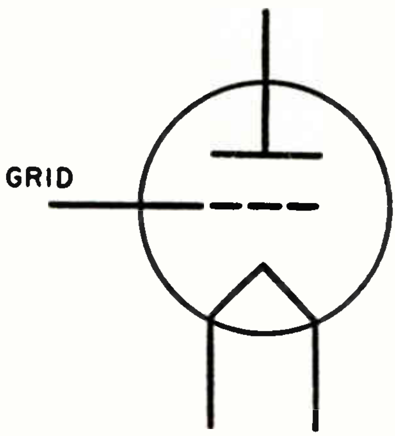


Fig. 30. The schematic symbol for a triode.

In a triode tube the filament is placed in the center of the tube. Around the filament, and close to it is a wire mesh; this is the grid of the tube. Placed some distance from the grid and around it is a round cylinder and this is the plate.

Because the grid is placed so close to the filament, a small voltage applied to the grid will have a large effect on the number of electrons that can flow from the filament to the plate of the tube. If the tube is connected into a circuit as shown in Fig. 31, you can see what will happen.

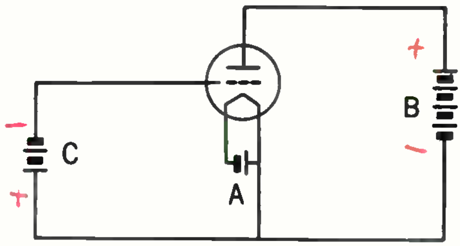


Fig. 31. A triode tube showing how three batteries are used to provide the necessary operating voltages.

The battery marked A is used to heat the filament of the tube. In early days of radio this battery was called the A battery. A small battery having a voltage of 3 or more volts is connected between the filament of the tube and the grid. This battery is labelled C on the diagram, and is called a C battery. Notice that the positive terminal of this battery is connected to the filament, and the negative terminal is connected to the grid. A battery having a somewhat higher voltage is connected between the plate of the tube and the filament. The positive terminal of this battery is connected to the plate and the negative terminal is connected to the filament. This battery is called a B battery.

Now let us see what happens in the tube. The filament of the tube is heated by the current from the A battery. This causes the filament to give off electrons and the electrons fly off into the space between the filament and the grid. However, many of the electrons are repelled by the negative charge on the grid of the tube due to the voltage of the C battery, and travel back to the filament. Some of the electrons manage to get through the grid and they are attracted by the positive potential applied to the plate of the tube by the B battery, and will travel over to the plate. The amount of current flowing from the filament to the plate of the tube can be controlled by the grid voltage. If we increase the negative voltage applied to the grid, the amount of current flowing through the tube will decrease, and if we reduce the negative voltage applied to the grid of the tube, the amount of current flowing from the filament of the tube to the plate will increase.

It is this ability of the grid to control the flow of current from the filament of the tube to the plate that makes the vacuum tube so useful in electronics. You will study vacuum tubes in detail in later lessons. For the present, you should remember how electrons flow from the filament of the tube to the plate of the tube and how the grid can control the flow of electrons through the tube. You should also remember the schematic symbols used to represent a diode (a two-element tube) and a triode (a three-element tube).

The schematic symbol of a triode tube with a cathode instead of a filament is shown in Fig. 32. In the triode tubes we have shown previously, the filament was heated by a battery and the filament gave off the electrons

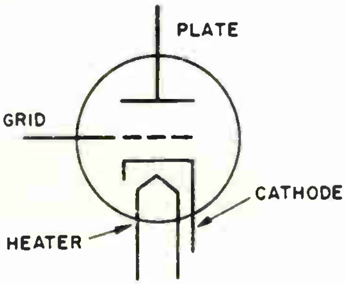


Fig. 32. Schematic symbol of a triode tube with a cathode and a heater.

that were used in the tube. More modern tubes have a cathode that is designed to give off the electrons. The cathode is a hollow round tube and it is coated with a special material that readily gives off electrons when it is heated. Inside of the cathode is a heater. The heater is heated by an external voltage applied to it and the heat from the heater radiates to the cathode and heats the cathode to a temperature where it will give off electrons.

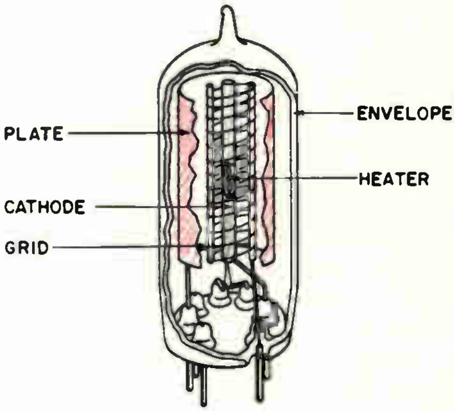


Fig. 33. Cut-away view of a typical vacuum tube showing its elements.

The cathode-type tube has replaced the filament-type tube in modern electronic equipment. The tubes used in modern radio and television receivers are all cathode-type tubes. A cut-away view of a modern tube is shown in Fig. 33. Filament-type tubes were used in portable receivers, but these have been replaced today by transistors.

TRANSISTORS

Transistors are made out of materials called semiconductors. Remember that a conductor is a material that will conduct or pass the flow of electric current. An insulator is a material that will not normally pass an electric current. A semiconductor is a material that falls midway between the two. It is neither a good conductor nor a good insulator.

Two materials, germanium and silicon are widely used in making transistors. Almost all the early transistors were germanium transistors, but now silicon transistors are about as numerous as germanium transistors. In the early days of semiconductors, manufacturing techniques had not been developed for the manufacture of silicon transistors. The few silicon transistors that were available were much more expensive than germanium transistors. However, today both types are widely available and there is very little difference between the price of the two.

A typical transistor is made up of three pieces of germanium or silicon as shown in Fig. 34. These three pieces are arranged as shown. Each piece of the germanium or silicon has been mixed with small quantities of another chemical. The pieces marked 1 and 3 have been mixed with the same chemical and the piece marked 2 has been mixed with small quantities of another chemical.

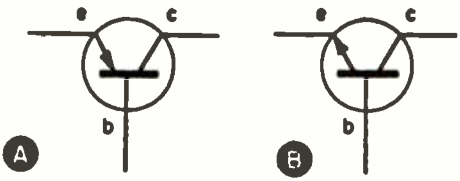


Fig. 35. Schematic symbols for two different types of three-element transistors. The emitter is marked e, the base b, and the collector c.

Since the transistor is made up of three pieces of material, it is often called a triode, just as the vacuum tube with the cathode, a grid and a plate is called a triode. The elements in a transistor are called the emitter, the base and the collector. The schematic symbols for transistors are shown in Fig. 35. The lead with the arrow on it and marked with the letter e, is the emitter, the long straight line marked b is the base and the other lead marked c is the collector. The two different types of symbols are for two different types of transistors. Their operation is somewhat different, but they can be used to accomplish the same thing.

The transistor is a comparatively new device compared to the vacuum tube. However, tremendous pro-

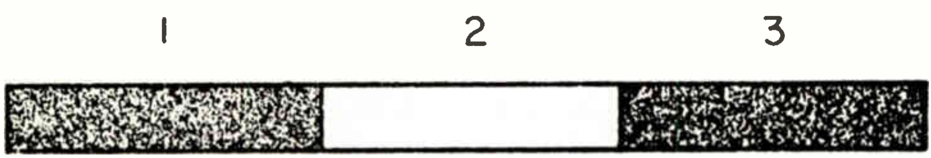


Fig. 34. A triode transistor made of three pieces of germanium. The germanium in the pieces marked 1 and 3 has been mixed with a small amount of one chemical; the germanium in the section marked 2 has been mixed with another.

gress has been made with the transistor in a relatively short space of time. Already, portable radio receivers using vacuum tubes have disappeared. All modern portable radio receivers use transistors. Also automobile radio receivers are now completely transistorized - they all use transistors. The only automobile radio you will find using vacuum tubes will be the automobile receiver designed for a car that is a number of years old.

Since transistors are making such important strides, it is extremely important that you learn all you can about them. For the present, simply remember the symbols used for them, and remember the names of the different elements in a transistor. Later, you will have several lessons devoted exclusively to transistors and you'll study transistor circuits in many of your more advanced lessons.

SUMMARY

In the preceding section you were introduced to many of the parts that you will study in detail in later lessons and will work with in your experimental kits. As we pointed out, it is important that you learn the schematic symbols for these parts as you go along, and also follow the simple circuits as you come to them. If you will do this as you go through your course you will find that schematic diagrams are easy to read, and you will soon be able to read fairly complicated diagrams without too much trouble.

There are a number of important points in this lesson that you should remember:

1. Conductors are used to carry currents from one part to another and insulators are used to keep the current from flowing where it is not wanted.
2. Batteries are made of groups of cells. If the cells in a battery are connected in series, that is the positive terminal of one connected to the negative terminal of the other, the voltages of the cells add so that the total battery voltage will be equal to the voltage of the cell times the number of cells in the battery.
3. When cells are connected so that the positive terminals of the cells are connected together and the negative terminals are connected together we say that the cells are connected in parallel.
4. Coils and transformers are widely used in electronics. A coil is simply made of a number of turns of wire. A transformer consists of two or more coils placed close together so that the magnetic lines of force produced when a current flows through one coil will cut the turns of the other coils. Remember the symbols used for air core coils and transformers and for iron core coils and transformers.
5. A capacitor is a device that can store an electric charge. It is made of two metal plates placed close together. Review the symbol used to represent a capacitor.
6. Resistors oppose the flow of current through them. Resistors having a resistance of only a few ohms up to resistances of over 1,000,000 ohms will be found in electronic equipment.
7. Two important types of vacuum tubes are the diode tube and the triode tube. A diode tube has two

elements, a triode tube has three elements.

8. Transistors are made of materials called semiconductors. Germanium and silicon are used in the manufacture of transistors. A transistor has three elements called an emitter, a base and a collector.

SELF-TEST QUESTIONS

(v) Name three materials that are good conductors.

(w) What metal is most widely used as an electrical conductor?

(x) If four 1-1/2 volt flashlight cells are connected in series to form a battery, what will the battery voltage be?

(y) Draw the schematic symbol of a 90 volt battery.

(z) Draw the schematic symbol for an air core coil.

(aa) Draw the schematic symbol for an iron core transformer.

(ab) What is the name of the device that can store an electric charge?

(ac) Draw the schematic symbol of a capacitor.

(ad) Name the three elements of a triode vacuum tube.

(ae) Name the three elements of a transistor.

Answers to Self-Test Questions

(a) Like charges repel, unlike charges attract.

(b) An atom is the smallest particle of an element that retains the original characteristics of the element. The atom will have a nucleus at its center and a number of electrons revolving around the nucleus.

(c) The nucleus and the electrons.

(d) The nucleus has a positive charge - the electrons have a negative charge.

(e) The copper atom will have a positive charge. If the atom loses one of its electrons, there will not be enough electrons to completely neutralize the positive charge on the nucleus and therefore the atom will have a positive charge.

(f) Electrons in adjacent atoms repel each other because all electrons have a negative charge and like charges repel.

(g) If you draw a diagram of the two atoms, you will soon see the answer to this question. The one atom that has the ten electrons will have two electrons in the first ring and eight electrons in the second ring. This is illustrated in Fig. 4 which shows the maximum number of electrons that can be in each ring. The atom that has eleven electrons will have two electrons in the first ring, eight electrons in the second ring, and the eleventh electron in the third ring. This atom will be quite unstable because the single electron in the third ring will not be held very closely to the nucleus. Indeed, this is the structure of the sodium atom. Sodium is a metal which in its pure state is so unstable that it must be kept submerged in oil. The atom with the ten electrons has the first two rings filled. You will remember that we said that atoms with the outer ring filled are very stable. This atom is indeed stable, it is the neon atom. Neon is called an inert gas by chemists. This means it is chemically inactive and will not combine with other elements.

(h) The negative terminal of a flashlight cell has a surplus of electrons. The positive terminal has a shortage. As a result, when a flashlight cell is connected into a circuit, the electrons leave the negative terminal, flow through the circuit and travel back to the positive terminal of the cell.

(i) When an electric circuit is completed, electrons start in motion instantaneously throughout the entire circuit. In a simple circuit such as shown in Figures 11 and 12, the number of electrons in motion is the same in all parts of the circuit.

(j) We measure electric currents in amperes. We usually abbreviate amperes, amps.

(k) The unit used to measure electromotive force is the volt.

(l) The unit used to measure the opposition to current flow in an electric circuit is the ohm.

(m) Increasing the voltage in an electric circuit will increase the current flowing.

(n) If we reduce the resistance in an electric circuit, the current flowing in the circuit will increase.

(o) The basic law of magnetism is, "like poles repel; unlike poles attract".

(p) The basic law of magnetism will give you the answer to this question. Since the pole of the magnet attracts the south pole of the compass, the pole must be an unlike pole, therefore it is a north pole.

(q) The magnetic lines of force leave the north pole of the magnet and travel through space to the south pole of the magnet.

(r) The magnetic lines of force around a current-carrying wire are strongest close to the wire. The further you get away from the wire, the weaker the magnetic lines of force will be.

(s) Placing an iron core inside of an electromagnet will increase the strength of the magnetic field.

(t) The flashlight bulb will light while the magnet is being placed inside of the coil. Once the magnet is inside of the coil and no longer moving, the flashlight bulb will no longer light.

(u) In the circuit shown in Fig. 20, when the switch is first closed, the magnetic field in L1 builds up slowly. This causes the number of magnetic lines of force to increase and the changing number of magnetic lines of force cutting L2 induces a voltage in L2. Once the magnetic field around L1 is built up to its full strength, the number of magnetic lines of force cutting L2 will no longer change and there will be no voltage induced in L2, therefore the flashlight bulb will no longer light.

(v) Copper, silver, aluminum, iron and most metals are good conductors.

(w) Copper.

(x) When the flashlight cells are connected in series the voltage will be equal to the cell voltage times the number of cells. 1-1/2 times 4 = 6 volts.

(y) See Fig. 22C, use the same symbol but write 90 v instead of 45 v.

(z) See Fig. 24A.

(aa) See Fig. 25B.

(ab) A capacitor.

(ac) See Fig. 26B.

(ad) Plate, grid and filament or cathode.

(ae) Emitter, base and collector.

Answering The Questions

On the last page of this lesson you will find ten questions. These questions are designed to help you learn the important points in this lesson. We do not want you to try to memorize the lesson or answer the questions from memory.

When you are ready to answer the questions, read over the first question carefully, make sure you understand the question and then mentally see if you can answer the question. Next, go to the section of the lesson where the answer is given, and read over that section of the lesson again. Make sure that you completely understand the answer to the question. Then, close the book and write out the answer. Do not copy the answer from the book, but rather try to write the answer in your own words. If you find that you have difficulty and cannot answer the question, it is an indication that you need to study some more.

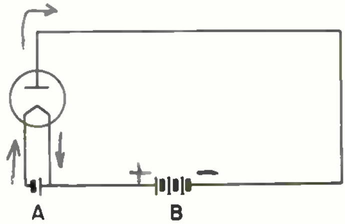
Many of the questions can be answered by a single word or by one or two words. Make your answers as brief and as direct as possible. Make sure that your answer actually answers the question asked. In some questions you will be asked to draw a schematic diagram or part of a diagram. Be sure you check these diagrams over carefully before you send in your answers for grading, because it is easy to make a mistake in drawing schematics.

Lesson Questions

Be sure to number your Answer Sheet B101.
Place your Student Number on every Answer Sheet.

Most students want to know their grades as soon as possible, so they mail their answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable. However, don't hold your answers too long; you may lose them. Don't hold answers to more than two sets of lessons at any time, or you may run out of lessons before new ones can reach you.

1. Two small balls of unknown material that are suspended near each other by threads are charged by being rubbed with a silk cloth. The balls then repel each other. Which one of the following statements will then be incorrect?
   (1) The balls both have a negative charge on them.
   (2) The balls both have a positive charge on them.
   (3) One ball has a positive charge, the other a negative charge.
2. Which does a negatively charged body have, a shortage or a surplus of electrons?
3. Draw a schematic diagram of a wire connected between the two terminals of a flashlight cell. Mark the polarity of the battery terminals and show by means of arrows the direction in which current will flow.
4. If we double the voltage applied to a circuit, what will happen to the current?
5. If you find that when you bring 2 magnetic poles together they attract each other, which one of the following statements is true?
   (1) The two poles must both be north poles.
   (2) The two poles must be south poles.
   (3) One pole must be a north pole and the other a south pole.
6. If a permanent magnet is held motionless inside a coil, will a voltage be induced in the coil?
7. How many flashlight cells would you have to connect in series to get 7.5 volts?
8. Will current flow from the filament to the plate in the circuit shown at right? Explain your answer. (Notice that the negative terminal of battery B is connected to the plate of the diode.)



9. Draw the schematic symbol of a modern triode tube and label the parts.
10. Draw the schematic symbol for a transistor and label each element with its full name.

HOW TO BUILD CONFIDENCE

Self-confidence--an active faith in your own power to accomplish whatever you try to do--is a personal asset which can do big things for you.

One thing which builds self-confidence is a successful experience. Each lesson completed with a passing grade is a successful experience which will build up confidence in you.

Little successes are contagious. Once you get a taste of success, you'll find yourself doing something successful every day. And before you realize it, your little successes will have built up to that big success you've been dreaming of. So--get the habit of success as fast as possible. Resolve to study every day, even if only for a few minutes.

Another confidence builder is a deep, firm faith in yourself--in your ability to get ahead. If you do believe in yourself and you are willing to back up this faith with good hard studying, you can safely leave the final result to itself. With complete confidence, you can look forward to an early success in electronics.

Act as if you could not possibly fail, and you will succeed!

J. F. Thompson

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-201

ACHIEVEMENT THROUGH ELECTRONICS

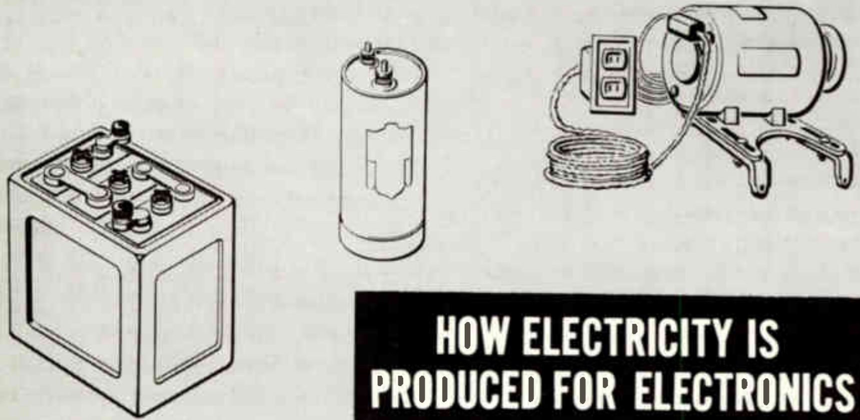
HOW ELECTRICITY IS PRODUCED FOR ELECTRONICS

B102

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

HOW ELECTRICITY IS PRODUCED FOR ELECTRONICS

B102



HOW ELECTRICITY IS PRODUCED FOR ELECTRONICS

Electronics is a term you will be meeting constantly from now on. Let us take time to see what we mean by the word "electronics." Originally the term electronics was applied only to devices using electronic tubes. However, in recent years the meaning has been broadened to include the whole field of electron behavior. Thus we can consider every application of electricity as part of the general field of electronics.

You have already seen many uses of electronic principles. You see them every day, for example, in your radio and television receivers.

The field of electronics is a growing field. The chances are that at this very minute while you are reading this lesson, engineers are working on new projects that will result in some new use of electronics. They may be working on some method of using electronics to improve on some process we are already using, or they may be working on something that is entirely new, something we have been unable to do before. You can be sure that we are going to see many new developments in electronics in the years to come.

In this lesson you are going to study some of the uses of electronics. We will cover only a few of the details, but you will learn enough to be able to understand these uses of electronics. In addition to learning something about these processes, you will learn more about the behavior of the electron.

The first use of electronics that you will study is in the power industry. Although this is often not considered part of the electronics industry, it is extremely important to electronics, because without economical power there would be very little use of electronics at all. This brief look at the power industry will help you in your study of other pieces of electronic equipment.

You will also learn more about electronics in radio. You will study several new components, and also you will learn how sound is sent through space by means of radio waves. This section on radio is important because the circuits you will study here are similar to ones you will find in radio, in television, and in other industrial applications.

You will also see how electronic principles are used in one branch of industry. Many industrial processes require large amounts of direct current.

It is much more convenient to generate and transmit alternating current than direct current. You will see how alternating current can be changed into direct current by means of electron tubes.

Finally, you will learn about pulse-type signals that are widely used in computers and television. You will learn the basic fundamentals of logic circuits that process these pulse signals in computers.

In studying the following sections of this lesson, it is important for you to understand the basic circuits and ideas presented. However, it is not necessary for you to remember all of the details of the various processes. As you complete a section of the lesson, answer the self-test questions at the end of the section. If you can answer these questions and remember the answers to them, then you should have no difficulty answering the questions at the end of the lesson. Between the self-test questions and the lesson questions, we will cover all of the important points in the lesson. Other details of the lesson will become more familiar to you as you go further in your course; these details will be covered again, and in many cases explained much more thoroughly than in this lesson.

If there are any basic circuits and ideas that you do not understand be sure to go over these points several times. If you need help be sure to take advantage of the NRI consultation service and write in requesting assistance. A thorough understanding of the basic fundamentals is absolutely essential; the more advanced lessons you will study later will be based on the fundamentals you will be learning in this and other basic lessons.

We have already mentioned that without economical electric power, there would be no electronics industry. In addition to economical power, we must have large amounts of power available.

The information on how electric power is generated is important to the radio-TV serviceman, the computer technician, the communications technician, and the industrial electronics technician. All may have occasion to service power-generating equipment. Even the radio-TV serviceman with only a small business may want to fix mobile equipment. There is more and more demand for people who can do repair work of this type, and in order to do such repairs one must be familiar with power-generating equipment.

There are two main sources of power in electronic equipment. One is batteries which you have already studied briefly. You will now look into batteries in more detail, and then go on to the study of generators to see how they operate and how the voltage generated is pictured.

Batteries

Batteries can be divided into two types: those containing primary cells and those containing secondary cells. Primary cells are cells that cannot be recharged. A flashlight cell is an example of this type. A secondary cell is a cell that can be recharged. The storage batteries used in automobiles are made up of secondary cells as are the rechargeable batteries used to operate portable television receivers designed for battery operation.

We mentioned previously that a battery is made up of two or more cells. But the word battery is also used to describe single cells, such as flashlight cells. Consequently, the word battery has come to mean anything from one cell up.

A basic knowledge of batteries is important to the electronics technician today because many of the devices he will encounter operate from batteries. For example, there are millions of portable radio receivers in use; almost all of these operate from batteries made up of cells like the flashlight cell. Portable television receivers which can be operated from either the power line or from a battery pack are becoming increasingly popular. These television receivers operate from small storage batteries that are similar to the storage batteries used in automobiles. Other types of batteries, such as the manganese battery and the mercury battery, are becoming increasingly important. In some ways these cells are similar to the flashlight cell, but they have a longer life. The chances are that they would replace the flashlight cell entirely except for the fact that they are more expensive.

A simple electric cell can be made by inserting two pieces of metal in an acid solution. The voltage produced by the cell will depend upon the metals used. For example, if one metal is zinc and the other copper, the cell voltage will be about 1.1 volts. If the copper electrode is replaced by a silver electrode, a voltage of about 1.5 volts will be produced. On the other hand, if magnesium is used as one electrode and gold as the other, a voltage of about 3.7 volts will be produced. Such a battery, while it has a relatively high voltage, isn't practical because of the cost of the gold and magnesium.

The cells we were speaking of contained acid in a liquid form. This type of cell has the disadvantage that the acid is easily spilled. A much more practical cell is the "dry cell", such as the flashlight cell, which we will now study in detail.

DRY CELLS

Dry cells are not really dry. The chemical mixture in the battery is actually quite moist, and when it becomes dry the battery is no longer usable. The name dry cell was given to these batteries because the chemical mixture was in the form of a paste rather than a liquid. The dry cell uses a carbon rod as the positive electrode and the zinc case as the negative electrode. The voltage of a cell of this type is 1-1/2 volts.

This cell is actually similar to the simple basic cells we mentioned earlier which contain two metals in an acid solution. In this case, one metal is zinc and the other is carbon. We do not often think of carbon as a metal, but actually it is midway between the metals and non-metals, and in some cases acts like a metal and in other cases acts like a nonmetal. When used in a dry cell it acts like a metal.

The acid is in the form of a paste made up of ammonium chloride, powdered carbon and manganese dioxide. The ammonium chloride is the acid, the other two chemicals are added to improve the performance of the cell.

The construction of one type of dry cell that can supply current for a much longer time than a flashlight cell is shown in Fig. 1. This cell is similar to a flashlight cell except it is larger and has screw-type terminals to which the leads are connected. The metal case and carbon rod are often provided with screw terminals as shown here. Sometimes a plug type of

Fig. 1. Construction of a dry cell.

Fig. 2. A square 1½-volt battery.

connector is provided instead. A cell of this type can provide a much higher current than a flashlight cell, because it has much larger electrodes. However, the output voltage of the cell is the same as that of the flashlight cell (1-1/2 volts). In the early days of radio, four cells of this type, connected in series, were used to provide 6 volts to operate the filaments of the tubes in many radios.

In addition to the round cell shown in Fig. 1, 1-1/2-volt cells are often made square. A square cell can often be fitted into a somewhat smaller place than a round cell, and therefore is particularly useful in portable equipment. A square cell is shown in Fig. 2. It would be more correct to call a square cell a battery because it is generally made up of four small round cells instead of one large cell. The negative terminals of all four cells are connected together and brought out to one common terminal which is the negative terminal of the battery. Similarly, the positive terminals of all four cells are connected together and brought to one common positive terminal.

As you learned earlier, this type of

connection is called a parallel connection and though there are four cells used, the output voltage from the battery is only 1-1/2 volts, the same as it would be if only one large cell were used. As we pointed out earlier, the advantage of this type of construction is that the four cells in parallel can supply more current than a single cell could alone. This is because the current the cell can supply depends primarily on the area of the positive and negative electrodes -- the carbon rod and the zinc case. Their area is greater when four cells are used in parallel than it would be if one cell of the same physical size were used. In addition, we pointed out that it is often more economical to use the four cells because the manufacturer may be making them in large quantities for other uses. This will bring the cost down so that it is more economical to use four of the smaller cells than it would be to use one large cell.

You already know that dry cells can be connected in series to provide more than the 1-1/2 volts available from the single cell. An example of this type of battery is shown in Fig. 3. This battery is designed

"B" BATTERY

Fig. 3. A 45-volt battery made of dry cells connected in series.

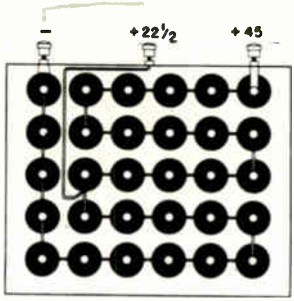


Fig. 4. How the dry cells in Fig. 3 are connected in series.

to provide a voltage of 45 volts and is called a "B" battery. This type of battery was widely used in the early days of radio to provide the voltage between the plate and filament of the tubes in radio receivers. Later, smaller versions of the battery were used in portable receivers. Today, the battery itself is not as important as the lesson we can learn from it about the voltages between different points. The following section of this lesson is extremely important, read it several times to be sure you understand it completely.

The battery shown in Fig. 3 is made up of thirty 1-1/2 volt cells connected in series as shown in Fig. 4. Notice that there are three terminals brought out of this battery. It is easy to see that the voltage between the two outside terminals should be 45 volts. There are thirty 1-1/2-volt cells, and 30 X 1-1/2 is 45. Now trace out the circuit between the negative terminal and the terminal marked +22-1/2 and you will see how we get this voltage. You will find that there are fifteen cells connected between these two terminals, and 15 X 1-1/2 is 22-1/2. Thus, if you need only 22-1/2 volts you would connect between the – terminal and the terminal marked +22-1/2.

Now look at the other half of the battery. What is the voltage between the terminals marked +22-1/2 and 45? By inspecting Fig. 4 you can see that there are fifteen cells connected between these two terminals. 15 X 1-1/2 is 22-1/2, and therefore there should be 22-1/2 volts between these two terminals. But which terminal is positive and which is negative? By looking at Fig. 4 again, you can find the answer to this question. Notice that the terminal marked +45 is connected to the positive terminal of one of the cells. The terminal marked +22-1/2 is connected to the negative terminal of the last cell in the group of fifteen cells connected between these two terminals. Therefore, this is the negative terminal and the +45 terminal is the positive terminal.

It may seem somewhat confusing at first that the terminal marked +22-1/2 can be both positive and negative. Let us see how this can be so. Starting with the negative terminal and looking toward the other two, you first see a group of fifteen cells and then a terminal. This terminal is positive compared to the negative terminal. We say it is positive with respect to the negative terminal. Then there is another group of fifteen cells and another terminal. This last terminal is even more positive with respect to the negative terminal. Now, if we started at the positive terminal and looked back through the battery we would see a group of fifteen cells and a terminal. This terminal is negative with respect to the positive terminal. We would then see an additional group of fifteen cells and another terminal, which is even more negative with respect to the positive terminal. We could, if we wished to do so,

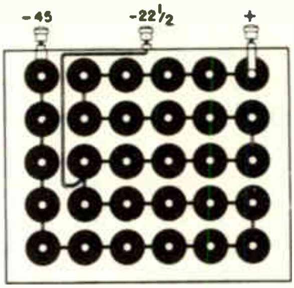


Fig. 5. The terminals of the 45-volt battery which are shown in Fig. 4 can be marked as shown.

mark the battery as in Fig. 5. Notice that this is the same battery as the one in Fig. 4; we have simply marked the terminals differently. The voltage between the two outside terminals is still 45 volts, and the voltage between either outside terminal and the center terminal is 22-1/2 volts. In Fig. 4, we have considered the voltage at the negative terminal as zero and marked the other two positive with respect to it. In Fig. 5, we have considered the positive terminal as zero volts and marked the other two negative with respect to it. We could go one step further and mark the center terminal zero and the one outside terminal −22-1/2 volts and the other +22-1/2 volts with respect to the center terminal.

You might wonder why there are all these different ways of marking battery terminals. The reason is that in electronic equipment one terminal of a battery is usually connected to a common or ground terminal in the equipment. Sometimes this terminal is the positive terminal of the battery, sometimes it is the negative terminal. It all depends on what the battery is to be used for. Usually the terminal voltages are marked with respect to the terminal that will be grounded in normal operation.

It is important to understand what is meant by a ground terminal or connection. In the early days of radio almost all radios were connected by a wire to a water pipe or to a pipe driven into the ground to improve reception. This was called a "ground" lead. One terminal of the "A" battery, the negative terminal of the "B" battery and the positive terminal of the "C" battery were all connected to the metal chassis on which the receiver was built, and the metal chassis was connected to the ground lead. Now the chassis in electronic equipment is called a chassis ground even though it may not be connected to an external ground connection at all. The negative side or terminal of the B supply is called B– ground or the common ground. In some equipment B– is connected to the chassis, but in other equipment it is not. When B– is connected to the chassis, we refer to both B– and the chassis as ground. When B– is not connected to the chassis, we refer to the chassis as a chassis ground and to B– as a floating ground. You will see these three expressions used frequently.

Remember that one terminal can be both positive and negative at the same time. In other words, one terminal might be positive when it is compared to another terminal, but negative when compared to a third. You will run into this situation over and over again.

Although a B battery is able to supply a much higher voltage than a single cell, the amount of current that can be taken from it is somewhat limited. If a B battery is used to supply a high current, it will soon be exhausted.

When you need a higher current than

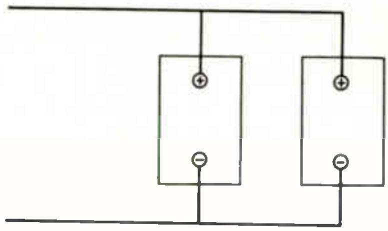


Fig. 6. Two batteries having the same voltage can be connected in parallel if a high current is required.

can be supplied by a single cell or battery, you can obtain it by connecting two or more batteries in parallel. When you connect two batteries in parallel, you connect the two negative terminals together and the two positive terminals together as shown in Fig. 6. This is the same type of connection that is used in the battery shown in Fig. 2, where the 1-1/2-volt battery was made from four cells connected in parallel. Of course, you can only do this if the batteries have the same voltage. When two similar batteries that are connected in parallel are connected to a circuit, each battery will supply approximately one-half of the current used in the circuit. This is more economical than taking the full current from one until it is exhausted and then taking the full current from another, because the batteries connected in parallel will usually last more than twice as long as a single battery.

There are two big disadvantages of the dry cell. It has a fairly short shelf life and a rather large cell is required to supply a moderate current. The shelf life of a cell is the length of time it can be kept after it is made before it deteriorates to such an extent that its life is affected appreciably. When we say a dry cell has a short life we mean that it cannot be kept too long before it is put into service, otherwise it will not last long.

MERCURY CELLS

A cell that overcomes the two main disadvantages of the dry cell is the mercury cell. The voltage supplied by one of these cells is about 1.35 volts or about 1.4 volts depending upon the materials used in the cell.

There are two different types of mercury cells in use. One is a flat cell that looks something like a button, and the other is a cylindrical cell that more closely resembles a standard flashlight cell. The advantage of the button-type cell is that several of them can be stacked inside of one container to form a battery. A typical battery made of three flat cells is shown in Fig. 7. The battery is slightly smaller than a standard flashlight cell, but produces a higher voltage, has a longer life and can supply more current than a flashlight cell.

Fig. 7. The mercury battery shown is slightly smaller than a standard flashlight cell, but produces a higher voltage, and has a longer life than the flashlight cell.

Mercury cells were originally developed for use by the Armed Forces, but are now found in many pieces of portable equipment manufactured for civilian use. Their small size and long life make them ideal for use in transistorized equipment. Since transistors use only a small amount of current, transistorized equipment powered by mercury batteries can be operated for a long time before it is necessary to replace the batteries. When it is necessary to replace a mercury battery, the old batteries must not be disposed of by burning; these batteries will explode if they are thrown into a fire and might cause considerable damage or injury to someone nearby.

One other important point to remember about the mercury cell is that the small terminal on the top of the cell is the negative terminal. The case is the positive terminal. This is the opposite of the standard flashlight cell where the button on the top of the cell is the positive terminal and the case is the negative terminal.

MANGANESE BATTERIES

The big disadvantage of the mercury battery is that it is quite expensive. This has limited its use somewhat in entertainment type equipment, such as portable radios, etc. A battery that is not as expensive, but has many of the desirable characteristics of the mercury cell, such as long life, is the manganese cell. This cell is also often called the alkaline-manganese-zinc cell.

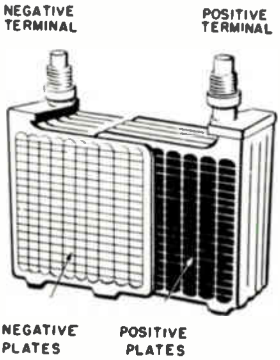
The manganese cell looks very much like a flashlight cell and is generally made in the same physical sizes as the small dry cells used for flashlights and portable radios. The voltage of a cell of this type is 1.5 volts so it can be readily substituted for a dry cell. A manganese cell can supply a certain value of current for much longer than the same size dry cell. Manganese cells are particularly useful in applications where the equipment is to be left on for a long time. Here their life will greatly exceed the life of a dry cell. However, in applications where equipment is operated intermittently for short periods, the manganese cell will outlast the dry cell, but the increase in performance may not be justified by the increased expense of the manganese cell.

LEAD-ACID BATTERIES

The storage battery used in automobiles is the best known example of a secondary cell. This type of cell has two groups of plates: one attached to the positive terminal and the other to the negative terminal. The plates are made of lead and fit together as shown in Fig. 8A. Between the plates are sheets of insulating material called separators, made either of porous wood or perforated wood or fiber glass. The separators prevent the plates from touching each other and destroying the cell. One of the sets of the plates is treated chemically to form an oxide of lead (a combination of lead and oxygen), and the two sets of plates with the separators between them are then placed in a container filled with a solution of sulphuric acid in water. The term "lead-acid battery" came about as a result of the use of lead plates in a solution of sulphuric acid.

The voltage in this type of cell is approximately 2 volts. Storage batteries used in modern automobiles are usually made of six cells connected in series so that the output voltage from the battery is 12 volts. Older automobiles used batteries in which three cells were connected in series to give an output voltage of 6 volts. A three-cell battery is shown in Fig. 8B.

To charge this type of battery, the battery is connected to a battery charger which simply applies a voltage slightly



(A)

(B)

Fig. 8. (A) shows the construction of a storage battery; (B) shows a three-cell battery.

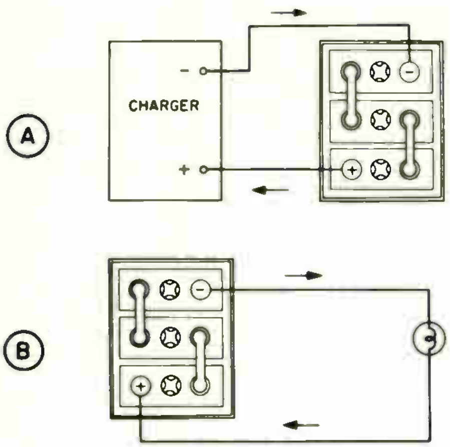


Fig. 9. (A) shows the direction current will flow when a storage battery is being charged; (B) shows the direction current will flow when the battery is supplying current.

higher than the battery voltage. The charger forces current through the battery as shown in Fig. 9A. This current causes a chemical change in the battery, and the electrical energy put in the battery is stored in it in the form of chemical energy. When the battery is connected to a circuit, the energy stored in the battery is released by the chemical action of the battery and current will flow in the circuit as shown in Fig. 9B.

Notice the direction of current in Figs. 9A and 9B. When the battery is being charged, the current is forced through the battery in the opposite direction to which it flows when the battery is supplying current. The storage battery can supply current for a much longer time than the average dry cell. When the storage battery is discharged and is no longer able to supply the current required by the circuit, the battery can be removed from the circuit and recharged by passing current through it in the opposite direction. Once the battery has been recharged, it can again be connected to the circuit and will supply current to the circuit.

In an automobile, the battery is usually connected to a generator. As long as the car is running at a reasonable speed, the generator is both charging the battery and supplying the current needed to operate the car. However, when the car is operating at a slow speed or when it is stopped, the generator is not turning fast enough to provide the electricity needed by the car and the battery supplies this energy. In the next section of this lesson you will study generators and you will see how they produce an electric current.

NICKEL-CADMIUM BATTERIES

There are two main disadvantages of the lead-acid storage battery. One disadvantage is that it has a somewhat limited life and the other is that it gives off hydrogen and oxygen when it is charged, and therefore the cell must be vented to allow these gases to escape. The hydrogen and oxygen given off from the battery come from the water in the cell and this water must be replaced periodically. As a result, the cell requires considerable maintenance.

A cell that overcomes this disadvantage is the nickel-cadmium cell. While the cell does give off gases when it is being charged, methods have been developed to take care of these gases so that the cell may be sealed. As a result, you do not have to add water or acid to the cell; the cell requires no maintenance other than to charge it when it becomes discharged.

The nickel-cadmium cell has an operating voltage of about 1.2 volts. While the cell cannot supply quite as much current as the same size lead-acid cell can supply, the fact that it will last almost indefinitely if it is cared for and that it can be sealed and requires no maintenance has made it ideally suited for use in operating portable electronic equipment. As more portable transistorized equipment is developed, it is likely that this type of cell will become more important to the electronics technician.

In charging both the lead-acid and the nickel-cadmium storage batteries or cells, the technician should follow the manufacturer's recommendations and avoid charging either type of cell at a higher rate. This is likely to cause excessive heat which can cause the plates in the cell to warp and touch. Once this happens, the cell is shorted and it is no longer usable.

SUMMARY

In this section of the lesson we have covered three important primary cells and two important secondary cells. We do not expect you to remember how these cells are made, but you should remember the voltage of each cell and remember their important characteristics. There may be occasions when you will want to substitute one type of cell for the other to obtain improved performance. In order to do this you must know the voltage supplied by each type of cell so that you can be sure that you will have the proper operating voltage for the equipment. You must also know something about their characteristics so that you will be able to select a suitable replacement that will result in improved performance.

The most important characteristic of the dry cell is its economy. Its disadvantages are its limited shelf life and its limited current capabilities.

The mercury cell has a much longer shelf life than a dry cell, and a given size mercury cell is capable of supplying a much higher current than a dry cell. That is to say, it can supply the same current as a dry cell for a much longer time. Mercury cells have a voltage of about 1.35 or 1.4 volts, depending upon the materials used in them. The disadvantage of the mercury cell is its cost.

The manganese cell can provide a given current for a much longer time than a dry cell can. Its shelf life is better than a dry cell, but not as good as a mercury cell. The output voltage of a manganese cell is about 1.5 volts.

The storage cell has the advantage that it can be recharged and used again. There are two important types of storage cells, the lead-acid cell and the nickel-cadmium cell. The lead-acid cell has a voltage of about 2 volts and the nickel-cadmium cell has a voltage of about 1.2 volts.

The advantage of the nickel-cadmium cell over the lead-acid cell is that it can be sealed and does not require the periodic maintenance that the lead-acid cell requires.

SELF-TEST QUESTIONS

(a) What is the output voltage of a dry cell?

(b) If fifteen dry cells are connected in series to form a battery, what would the output voltage be?

(c) If eight dry cells are connected in parallel to form a battery, what would the output voltage be?

(d) If six dry cells are connected in series to form a 9-volt battery and the negative terminal is marked with a minus sign and the positive terminal marked +9V, how would you expect a center tap terminal connected between the third and fourth cells to be marked?

(This is a difficult question - look at Figs. 4 and 5 before you make up your mind as to what the answer should be.)

(e) What are the two main advantages of the mercury cell over the dry cell?

(f) What is the output voltage of a mercury cell?

(g) What is the output voltage of a manganese cell?

(h) Since the manganese cell is a better cell than the dry cell, why hasn't it completely replaced the dry cell?

(i) Name two types of secondary cells.

(j) Which type of secondary cell provides the longest life, and the most maintenance-free performance?

Generators

Although batteries are very useful, their ability to supply large amounts of power is limited. If a battery, even a lead-acid storage battery, is called upon to supply large amounts of current, it will soon be exhausted and must be removed from the circuit and recharged. Even if storage batteries could supply the large amounts of electricity consumed daily by the average large city, we would still have to have some way of recharging the batteries. Thus, we would have a need for a device other than a battery that is capable of supplying large amounts of electricity.

You already know that if the magnetic field cutting through a coil is varied, a voltage will be induced in the coil. If the magnetic field is made to vary rapidly enough, the voltage will be induced continuously in the coil; this voltage can be connected to an outside circuit and current will flow through the circuit. This is the principle that is used in electric power generators.

Before studying generators, let us learn about two important kinds of current you will have to deal with: direct current and alternating current.

DIRECT CURRENT AND ALTERNATING CURRENT

The current supplied by a battery always flows from the negative terminal of the battery through the external circuit and back to the positive terminal of the battery. The current always flows in one direction. We call this kind of current ***direct current.*** We usually abbreviate this ***dc.***

The voltage supplied by a battery that causes a direct current to flow is referred to as a ***dc voltage.*** The expression dc voltage means that the voltage causes a direct current to flow. This means that the polarity does not change. In other words, on a battery, one terminal is always the negative terminal and the other terminal is always the positive terminal. We say that the terminals always have the same polarity -- the polarity of the terminals does not change.

A dc generator is a device that generates a direct current. In other words, the current coming from the generator always flows in the same direction. This means that the terminals of the generator must always have the same polarity; one terminal will always be the negative terminal and the other terminal will always be the positive terminal.

Besides direct current, there is another important type of current with which you must be familiar. This type of current is called ***alternating current*** or simply ***ac.*** Alternating current differs from direct current in that its direction is continually changing. The current flows first in one direction and then in the opposite direction. This reversal of current occurs many times a second. The current supplied to your home is alternating current. The voltage that produces an alternating current is called an ***ac voltage.*** In order to produce a flow of alternating current the polarity of the terminals of the device causing the current flow must be continually reversing. In other words, when the current is flowing in one direction, one terminal must be negative and the other positive. When the current flows in the opposite direction, the polarity of the terminals must reverse so that the terminal that was the negative terminal must become the positive terminal and the terminal that was the positive terminal must become the negative terminal.

While alternating current cannot be used to operate such electronic devices as tubes and transistors, it has many useful applications; indeed our modern industries depend upon large amounts of alternating current being readily available.

Now we will go ahead and learn how a simple generator operates and at the same time learn more about alternating current and ac voltage.

A SIMPLE GENERATOR

A simple generator capable of generating electric power can be built as shown in Fig. 10. This generator consists of a single-turn coil placed between the poles of a magnet. As the coil is rotated, it will cut through the magnetic lines of force

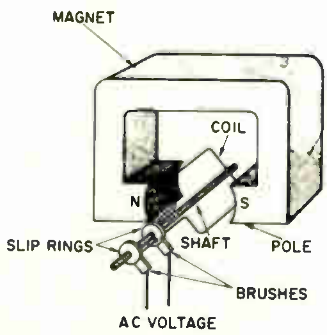


Fig. 10. A simple ac generator.

flowing from the north pole to the south pole of the magnet and a voltage will be induced in the coil. The amount of voltage will depend upon the number of magnetic lines being cut by the coil as it rotates. This in turn will depend upon the strength of the magnetic field and upon the speed at which the coil is rotated.

First, let us consider the voltage that will be produced by a generator of this type. When the coil is in the position shown in Fig. 11A, the movement of the coil is parallel to the lines of force flowing from the north pole to the south pole of the magnet. They are moving along the lines of force and are not cutting through any of the lines of force. You will remember that in order to induce a voltage in a coil, the turns of the coil must be cut by magnetic lines of force. In Fig. 11A, the coil is moving along the lines of force and not cutting through them. There will be no voltage induced in the coil when the coil is in this position.

When the coil moves counterclockwise to the position shown in Fig. 11B, it will still be moving almost parallel to the lines of force. However, it is moving at a small angle to these lines of force and therefore it will cut through some of them and there will be some voltage induced in the coil. When the coil moves down to the position shown in Fig. 11C, it will be cutting more lines of force because it is moving at a sharper angle to them and a somewhat higher voltage will be induced in the coil. Finally when it reaches the position shown in Fig. 11D, it will be moving directly perpendicular to the lines of force and will be cutting through them at maximum speed, and the voltage induced in the coil will reach its highest

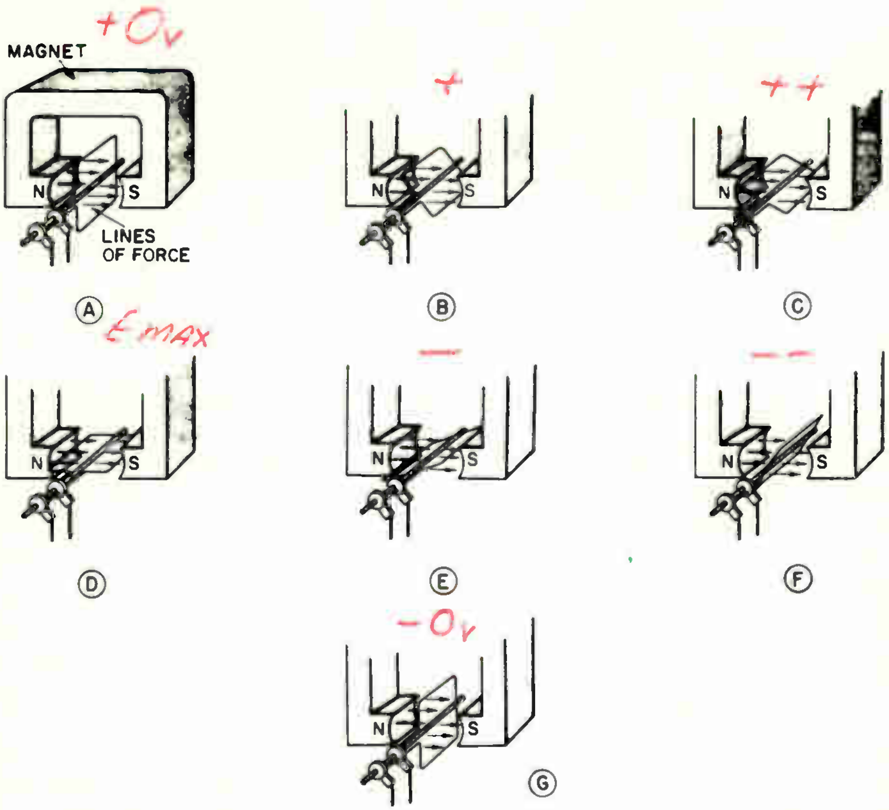


Fig. 11. In the illustration, the coil is rotating counterclockwise. The voltage produced by this generator depends upon the movement of the coil in relation to the magnetic lines of force.

value. As the coil moves down to the positions shown in Figs. 11E and 11F, it will be cutting fewer and fewer lines of force until it finally reaches the position shown in Fig. 11G. Once again the coil will be running parallel to the lines of force and no voltage will be induced in the coil.

HOW VOLTAGES ARE PICTURED

You will remember that when we discussed the batteries shown in Figs. 4 and 5 we said that one terminal of the battery could be considered as zero voltage and the voltage at the other terminals marked in reference to this terminal. You can do the same thing with a generator. You can use one lead as the ground or common lead and measure the voltage at the other lead as either positive or negative with respect to the common lead.

If we assume that one lead is a common or ground lead, we can conveniently represent the voltage at the other lead by a graph. A graph is merely a simple way of presenting information in the form of a picture. In Fig. 12A, you will notice a

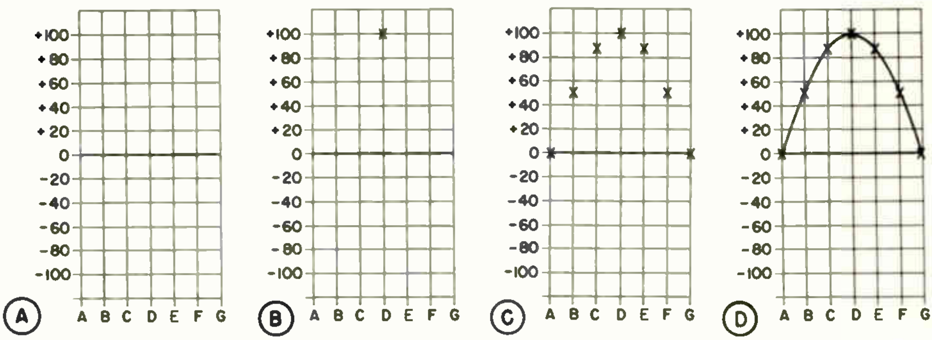


Fig. 12. Construction of a graph of the voltage produced by the generator in Fig. 11.

horizontal line which is marked zero running through the center of the graph. This is the zero voltage line which represents the voltage of the ground or common lead. The horizontal lines above this line represent positive voltages, and the lines below the zero line represent negative voltages. The vertical lines represent the positions .of the coil shown in Fig. 11.

Let us assume that at the instant the coil is in the position shown in Fig. 11D, the voltage generated is 100 volts. If the voltage is 100 volts positive with respect to the common terminal, we would place a mark (X) on the graph at the point where the +100 volt line crosses the vertical line running through D, as shown in Fig. 12B. Similarly the voltages that are present at the remaining points would be marked on the graph. This would look like Fig. 12C. The only step left is to draw a smooth curve joining all these points as shown in Fig. 12D. This curve represents the voltage generated by the generator through one half turn.

When the coil is rotated through the remaining half turn, the polarity of the voltage produced will be reversed. In other words, the voltage will now be negative with respect to the ground terminal because the coil will now be cutting through the magnetic lines of force in the opposite direction. If we complete the

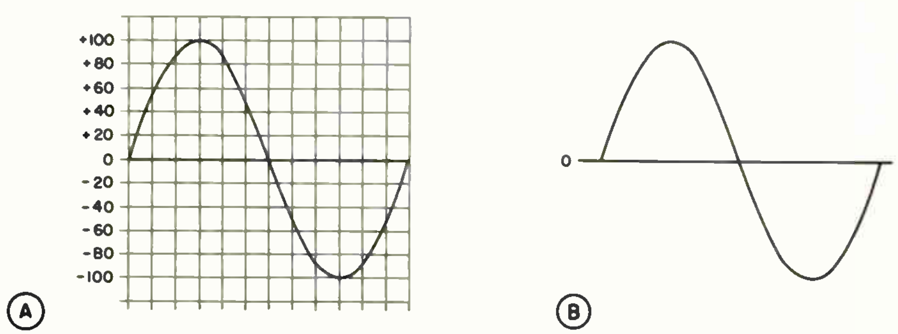


Fig. 13. The appearance of the output voltage produced by the generator as it travels through one complete turn.

drawing to show what the voltage will look like during the other half turn, the picture would look like Fig. 13A. For convenience, if we leave the horizontal and vertical lines of the graph off, we can get a better look at the shape or appearance of the output voltage as in Fig. 13B. This is called a waveform.

This is how the ac voltage supplied by the power company may be represented. It is called a sine (pronounced sign) wave.

The voltage represented by one complete turn of the coil in the magnetic field is called a cycle. The power supplied by most power companies in this country is 60-cycle power. When we say 60-cycle, we mean 60 cycles per second. In other words, the voltage goes through 60 cycles each second, like the one shown in Fig. 13. This is called the "frequency" of the ac voltage. To make the generator produce this type of voltage, it would have to be turned at a speed of 60 revolutions per second, which is 3600 revolutions per minute. The part of the cycle above the line is called a half cycle, and it is referred to as the positive half cycle. The part of the cycle below the line is also called a half cycle, and it is referred to as the negative half cycle.

You will also see the term Hertz, abbreviated Hz, used to designate the frequency of an ac signal. Hertz and cycles per second mean exactly the same thing and you will see these expressions used interchangeably. Hertz is the preferred term, but you will see cycles per second used almost as often. Just remember when you see the expression 60 Hz signal, we are referring to an alternating voltage or current that completes 60 cycles per second.

The voltage generated by this generator with only a single-turn coil would be extremely low, even with a very strong magnetic field. However, we can obtain a higher voltage simply by putting more turns of wire on the coil. Ten times as much voltage would be induced in a ten-turn coil as in a single-turn coil. If one-tenth of a volt is induced in a single-turn coil, one-tenth of a volt will be induced in each turn of the ten-turn coil. These voltages will be induced in series with each other so that the total voltage available at the output terminals of the coil will be 1 volt. By putting 100 turns on the coil, we could get 10 volts; by putting 1000 turns on the coil, we could get 100 volts.

Thus, by putting the proper number of turns on the coils of a generator we can generate any required voltage.

DC GENERATORS

Instead of using a permanent magnet such as the one shown in Fig. 10, a practical generator would use an electromagnet to supply the magnetic field. The current required to operate this electromagnet can be obtained either from the generator itself or from another generator. However, it must be dc. The voltage generated by the generator we have discussed is ac. Now let us look into the generator further and see how we can obtain dc instead of ac.

In an ac generator such as the one described, the ends of the coils are connected to slip-rings. By means of contacts riding on the slip-rings (which are called brushes), we could take the ac voltage off the generator. However, if, instead of using two slip-rings, we use a

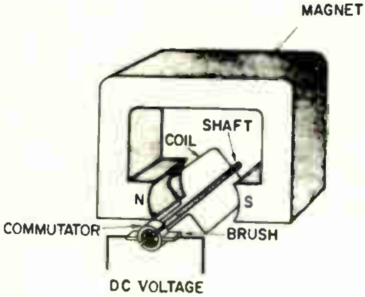


Fig. 14. A simple dc generator.

commutator, as shown in Fig. 14, we can obtain dc. A commutator is similar to a slip-ring except that it is split in half and the two halves are insulated from each other. The brushes are placed so that when the coil is going through the first half revolution, one brush will be connected to one section of the commutator and the other brush will be connected to

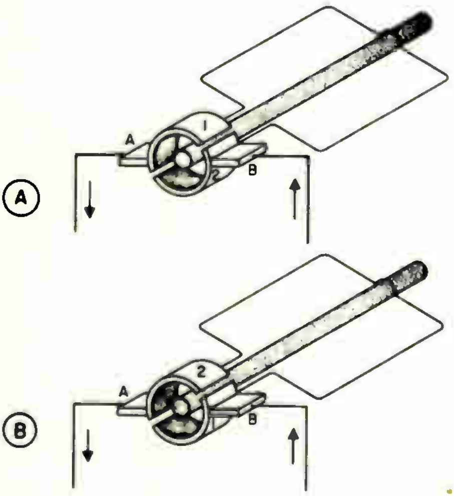


Fig. 15. In (A), section 1 of the commutator is negative, and section 2 is positive. In (B), section 1 is positive, and section 2 is negative. This means that brush A is always connected to the negative section of the commutator and brush B to the positive section, and current flows in only one direction.

the other section. When a coil starts to go through the second half revolution, the brushes will be connected to the opposite section; the one that was connected to the first section will then be connected to the second section, and the one that was connected to the second section will then be connected to the first section and current will always flow through the external circuit in the same direction, as shown in Fig. 15. As a result, at the output we will have a voltage like that shown in Fig. 16.

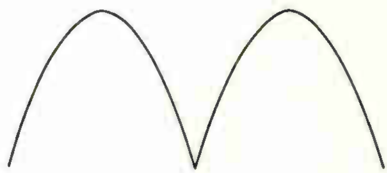


Fig. 16. Output of a simple dc generator.

Here we have essentially the same wave shape as we had in Fig. 13, except that the half of the cycle that was previously negative and drawn below the line has now been reversed by the automatic reversing of the connections to the coil performed by the commutator and brushes.

A voltage like that shown in Fig. 16 is called pulsating dc. It is dc inasmuch as the current that flows as a result of this voltage will always flow in one direction. However, since the voltage varies, the current varies; it will flow in pulses and actually drop to zero twice through each revolution of the coil. This can be used in some applications but it is not what we call pure dc, like the dc supplied by a battery.

As in the case of the ac generator, the output voltage from a dc generator with only a single turn would be extremely

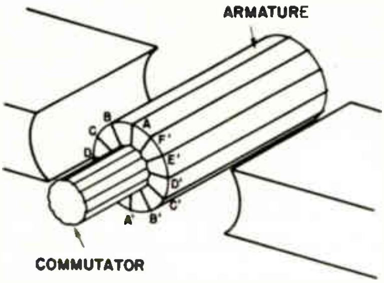


Fig. 17. The armature of a dc generator with 6 series-connected coils.

low – too low to be usable. However, by winding several turns on the coil, a higher voltage can be obtained. This will still provide a pulsating type of dc like that shown in Fig. 16.

A better arrangement is shown in Fig. 17. Here, there are a number of coils on an iron form called an "armature." The coils are in different positions around the armature. In the position shown, coil A-A′ will not cut any lines of force and therefore there will be no voltage induced in it. The coil B-B′, however, is cutting a few lines of force and some voltage will be induced in it. Coil C-C′ is cutting still more lines of force and a somewhat higher voltage will be induced, and coil D-D′ is cutting directly across the lines of force and the maximum voltage will be induced in it. These coils are all connected in series and brought out to connections on the commutator. Two brushes are used on the generator, and the output voltage from this type of machine will be nearly constant. This is because there will always be one coil either in or close to each position shown in Fig. 17. The voltage produced by the generator will be the voltage of all the coils in series. While some of the coils are producing very little voltage, the coils near the position D-D′ will produce considerable voltage. The commutator used with this type of generator would have twelve sections instead of two, as is the case in the generator with only one coil. There will be some fluctuation in the voltage so that the output voltage will look like Fig. 18 as the generator goes through one revolution. Notice that with this type of generator, instead of two big half cycles, an almost constant value of dc is obtained.

Fig. 18. Only a small ripple will be noticeable in the output of a dc generator having 6 coils and 12 commutator sections.

AC VALUES

Now look back at the ac cycle shown in Fig. 13. Notice that at the start of the cycle the voltage is 0. When the coil has rotated through one quarter of a turn, the voltage has built up to a maximum value; when the coil has gone through one-half a turn, the voltage is back to zero again. At three-quarters of a turn, the voltage has reached a maximum value with the opposite polarity and finally at the end of the cycle, the voltage is back to zero. The end of this cycle is actually the start of the next cycle. The voltage reaches a maximum value twice in each cycle, and it drops to zero twice in each cycle.

Let us see how we measure ac. When we talk about a dc current and we say that the current flowing in the circuit is 1 amp, we mean that a certain number of electrons are flowing past a given point in the circuit. This same number of electrons continues to flow as long as the current in the circuit is 1 amp. However, what happens when we're dealing with ac? Since the voltage reaches a maximum twice each cycle and falls to zero twice each cycle, the current must also have two maximums in each cycle and fall to zero twice each cycle. Therefore, the number of electrons flowing in the circuit is not constant; in fact it is continually changing as the ac voltage goes through its cycle.

To overcome this difficulty, we measure ac in terms of equivalent dc. When we say that the ac current flowing in the circuit is 1 amp, we mean that the current is the equivalent of 1 amp of dc. In other words, if a dc current of 1 amp flows through a heating element such as found in an electric iron or toaster, a certain amount of heat will be produced. When the ac current flowing through the same heating element produces the same amount of heat, we say that the ac current is 1 amp.

The same system is used to measure ac voltage. If a dc voltage of 100 volts is required to force the current of 1 ampere through a circuit made up of a resistance, the ac voltage that will force an ac current of 1 amp through the same resistance is said to be 100 volts. This is called the ***effective*** or ***rms voltage*** and it will cause an ac current to flow that will produce the same heating effect as the equivalent amount of direct current.

From looking at an ac voltage cycle, it is quite obvious that the voltage must be greater than the effective value during part of the cycle and less than the effective value during the remainder of the cycle. The maximum value that the ac voltage reaches during a half cycle is called the ***peak voltage***. The peak voltage is approximately 1.4 times the effective voltage. Each peak is 1.4 times the effective voltage, and the voltage between

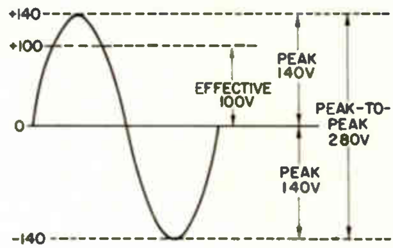


Fig. 19. AC waveform showing effective (rms), peak and peak-to-peak voltages.

the two peaks will be 2.8 times the effective voltage. This voltage is called the ***peak-to-peak voltage***. Fig. 19 shows an ac voltage of 100 volts. The peak value is 140 volts, and the peak-to-peak value 280 volts. Study this figure to be sure you understand what the terms mean.

It's important to remember when we are talking about ac voltages that unless we specifically refer to the peak or peak-to-peak voltages, we are talking about effective voltages; in other words, the ac voltages that will produce the same effect as the equivalent dc voltage. Also remember that the peak or greatest voltage reached during a half cycle will be 1.4 times as great as the effective voltage, and that the peak-to-peak voltage will be 2.8 times the effective voltage. You will run into the terms "peak voltage" and "peak-to-peak voltage" as well as the expression "effective voltage" many times during your career in electronics. Knowing the peak value of an ac voltage is often very important.

A good example of how high the peak voltage in a circuit may be is the voltage supplied in most homes in this country. The average power company supplies a voltage somewhere between 115 and 120 volts, 60 Hertz ac for lighting and general domestic uses. This is the effective value of the voltage. Let us assume that the voltage supplied to your home is exactly 120 volts. What is the peak voltage reached during each half cycle? It will be 1.4 X 120, which is 168 volts. In other words, twice during each cycle the actual voltage between the two power leads connected to each electric light and appliance in your home reaches a value of 168 volts. Twice during each cycle the voltage also drops to zero.

The net effect of this ac voltage is the same as supplying a dc voltage of 120 volts to the electric lights. The peak-to-peak voltage supplied to your home will be 2.8 X 120 volts or 336 volts!

THE IMPORTANCE OF AC

You may wonder why we have gone into so much detail in describing alternating current. Alternating current is important not only because it is the type of power supplied by the power companies, but also because ac signals are used throughout the whole electronics industry. The sound that comes from the loudspeaker in your radio or your television receiver and the sound from your telephone are produced by ac signals having a frequency not too much higher than the power line frequency. The radio waves that travel through space are actually ac signals of a much higher frequency. Signals used in many industrial applications are ac signals.

SUMMARY

We have covered a great deal of material in the preceding sections. There are several important things that you should remember. First, remember what an ac cycle looks like. Also remember that this ac voltage is called a sine wave.

Remember that when we speak of ac voltage and current we are speaking of the voltage and current that will produce the same effect as the equivalent values of dc. Remember that the peak value of the ac cycle is 1.4 times the effective value, and the peak-to-peak value is 2.8 times the effective value.

SELF-TEST QUESTIONS

(k) What are the two kinds of current that the electronic technician must know about?

(l) What do we call the voltage that produces a current that flows in only one direction?

(m) What is the name given to a voltage that causes current to flow first in one direction and then in the other direction many times in a second?

(n) What is the name given to the ac voltage wave supplied by the power company?

(o) What is a cycle?

(p) What do we mean by 60 cycles per second (60 Hertz)?

(q) What is the name given to the device on a dc generator that is used to reverse the connections of the coil to produce dc instead of ac?

(r) What type of current is produced by a simple dc generator such as the one shown in Fig. 14?

(s) What do we mean when we say that the ac current is 1 amp?

(t) If we say the ac voltage is 100 volts, do we mean the effective or the peak value of the ac voltage is 100 volts?

(u) If the rms voltage is 200 volts, what is the peak value of the voltage?

Electronics in Communications

Radio waves are sine waves like the one shown in Fig. 13. They are similar to the 60 Hertz ac power waves. The only difference is in the frequency of radio waves. The lowest frequency radio waves have a frequency of about 15,000 Hertz. We usually use kilohertz instead of Hertz when speaking of radio waves around this frequency. -- one kilohertz equals one thousand Hertz. Therefore, 15,000 Hertz is equal to 15 kilohertz (kHz). This is a very low frequency radio signal; there are only a few very high-powered stations operating on such low frequencies. They are used mostly by government services for world-wide communications.

The standard radio broadcast stations operate on frequencies between 550 kHz (550,000 Hertz) and approximately 1,700 kHz (1,700,000 Hertz). This group of frequencies is called a band; in this case, the broadcast band. The word band when it is used this way means nothing other than a group of frequencies.

Needless to say, ac signals of such high frequencies cannot be generated by means of mechanical generators such as those used to produce 60-cycle power voltages and current. These ac signals must be generated by electronic devices such as vacuum tubes and transistors.

We mentioned that the frequency of radio signals is often expressed in kilohertz rather than in Hertz. A kilohertz is equal to 1,000 Hertz. However, many broadcast services operate on such high frequencies that even the kilohertz is not large enough to provide a convenient method of expressing their frequency; so in addition to the kilohertz we use the megahertz. The megahertz is 1,000,000 Hertz and it is abbreviated MHz. There are 1,000 kilohertz in a megahertz. The standard broadcast band is 550 kHz to 1700 kHz. This can also be expressed as .55 MHz and 1.7 MHz. TV stations operate up in the megahertz region. Channel 2, the lowest TV channel, occupies the band of frequencies from 54 MHz to 60 MHz. Channel 13 occupies the band of frequencies from 210 MHz to 216 MHz. These channels are referred to as the vhf (very high frequency) TV channels. The FM broadcast band, which is located between channels 6 and 7, occupies the band of frequencies from 88 to 108 MHz.

TV channels 14 through 83 are referred as the uhf (ultra high frequency) channels. Channel 14 occupies the band of frequencies from 470 MHz to 476 MHz and Channel 83 occupies the band of frequencies from 884 MHz to 890 MHz.

Now let us look at some parts of a radio system. We will briefly describe a radio transmitter and a receiver. Let us look at the transmitter first.

THE RADIO TRANSMITTER

Radio transmitters are actually made up of a number of sections, each designed to do a specific job. Let's discuss some of the more important sections of a typical transmitter.

The sound signal to be transmitted starts in the microphone. There are a number of different types of microphones in use. We will look at one of the simpler types. Before we discuss the microphone, let's learn a little about sound.

Sound is a vibration set up in air or some other medium. The key you strike on a piano is connected to a hammer, which strikes a string that is tightly stretched on a frame. The string begins to vibrate and sets the air surrounding it into vibration. The frequency at which the string vibrates will determine the tone that you hear. Similarly, when we speak, the speech muscles in our throat set the air in our throat into vibration. This vibration is projected from our mouth and nose, and the vibration travels through the air. However, instead of producing a single frequency, our voices are actually quite complex and may produce vibrations of many frequencies, causing the different tones by which we can distinguish the voice of one person from that of another.

Microphones. A simple microphone is shown in Fig. 20. This microphone consists of a coil and a magnet and a flat metal disc which is called a diaphragm. Fastened to the diaphragm is a small light coil form on which a coil is wound. The coil is placed between the poles of a permanent magnet. When you speak in front of the microphone, your voice sets the air into vibration. The vibrating air will cause the diaphragm to vibrate. Since

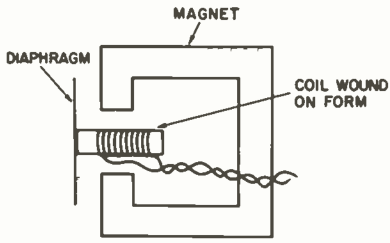


Fig. 20. A simple dynamic microphone.

the diaphragm is connected to the coil, the coil will vibrate between the poles of the magnet. You already know what will happen when a coil is moved in a magnetic field; the turns of wire on the coil will cut through the magnetic lines of force, and voltage will be induced in the coil. If this coil is connected to an external circuit, there will be a current flow through the circuit. The frequency of the current will depend upon the frequency at which the diaphragm and coil vibrate. This in turn will depend upon the frequency at which the air is vibrating, which in turn depends upon the vibrations set up by the vocal muscles in the throat of the person speaking in front of the microphone.

This electrical signal produced by the microphone is called an audio signal or audio voltage. The word audio is used to designate electrical signals in the frequency range of sound. An audio voltage or signal is the electrical equivalent of sound.

An Audio Amplifier. In a typical transmitter the output of the microphone is fed to an audio amplifier. The audio voltage is fed to a vacuum tube or a transistor that will amplify the signal so that a much stronger signal will appear in the output circuit of the tube or transistor. The amplified signal will be just like the signal generated by the microphone, but it will be stronger. A tube or transistor along with its associated parts is called a stage. A stage used to amplify an audio signal is called an audio stage.

In a typical transmitter, the audio signal produced by the microphone will be amplified by a number of stages in order to build up the strength of the signal until it is hundreds of times stronger than the original signal produced by the microphone. The signal is first fed to one stage where it is amplified and then on to a second stage where it is amplified still further and on to a following stage and so on until its strength has been built up to the desired level. The amplified signal, however, will have exactly the same frequency and characteristics as the original signal produced by the microphone.

The microphone and the audio amplifiers in a transmitter are called the audio section. This is often abbreviated af (audio frequency) section. Another important section of the transmitter is the radio frequency section -- this is abbreviated rf section.

The RF Section. The rf section of a transmitter is made up of a number of separate stages. The first stage is the stage that actually generates the radio frequency signal. This stage is called the oscillator stage. The oscillator stage is carefully designed to produce a signal of the desired frequency. The signal from the oscillator stage is then amplified by additional stages which are called rf power amplifier stages. These stages build up the strength of the radio frequency signals generated by the oscillator so that the signals will be strong enough to travel through space from the transmitter to the receiver.

In one of the radio frequency stages of the transmitter, the audio signal is superimposed on the radio frequency signal. This is called modulation. The audio signal is the modulation signal and the rf signal is the modulated signal.

In a radio system of this type, the modulation signal varies the amplitude or strength of the rf signal. This type of modulation is called amplitude modulation and is abbreviated AM. In some transmitters the modulated signal is then fed directly to the antenna, but in others it is amplified further and then fed to the antenna.

The modulated rf signal from the transmitter is fed through a cable or wire, called a transmission line, to the antenna. The transmission line is something like the power lines that are used to bring the power from the power-generating station to your home. They simply carry the power from the transmitter to the antenna.

The antenna at a radio station is simply a length of wire or a tower to which the transmission line is connected. When the radio frequency signal is fed to the

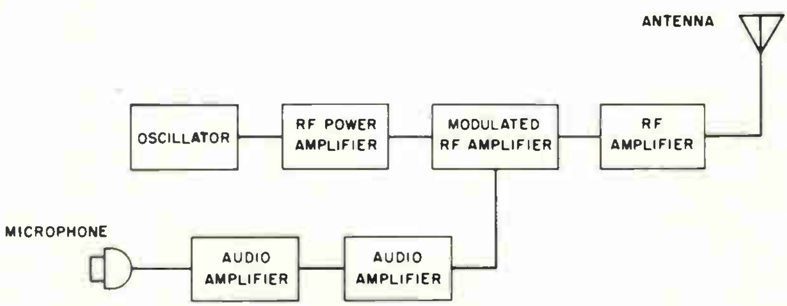


Fig. 21. A block diagram of a simple transmitter.

antenna, the rf signal sets up a current in the antenna which produces a magnetic field and an electric field surrounding the antenna. These fields travel out in space from the antenna and carry the signal from the transmitting antenna to the receiving antenna.

A simplified diagram of a radio transmitting system is shown in Fig. 21. This type of diagram is called a block diagram. It is a simpler way of representing the various stages or sections of a piece of electronic equipment than showing the complete schematic diagram. We will use this type of diagram frequently because it enables us to give an overall picture of the various stages of a piece of electronic equipment without going into all the details of the circuitry. You will find later that the circuits used in the various stages follow certain basic patterns. In other words, there is little difference in the circuits used in the individual stages; the difference is the manner in which the stages are used.

Notice that the sound signal is generated by a microphone and fed to the audio amplifier stages. At this same time the rf signal is generated by an oscillator and amplified by rf power amplifier stages. In the third rf stage, the rf signal is modulated by the audio signal and the modulated signal is then amplified by the last stage in the transmitter. This stage is called the final amplifier or simply the final because it is the last rf amplifier stage in the transmitter. The signal from the final is then fed through a transmission line to the transmitting antenna.

This brief description of Fig. 21 is all you need to know about radio transmitters at this time. However, you should understand in general terms what is done in a transmitter; it will be helpful to you, even though you may never work on a radio transmitter. Notice the symbols used in Fig. 21 for the microphone and the antenna. Remember these symbols; you will see them frequently in future lessons.

THE RECEIVER

The job that the receiver has to perform is exactly the opposite of the job the transmitter has to perform. The transmitter must take the sound and convert it to an electrical signal, which is the audio signal, and then superimpose the audio signal on an rf carrier signal. The receiver must take the rf carrier and remove the audio signal from it and then convert the audio signal back into sound.

In spite of the fact that the receiver must perform the opposite tasks from those performed by the transmitter, there are many similarities between a transmitter and a receiver. A transmitter has rf amplifiers, so do many receivers. The transmitter has audio amplifiers, so do receivers. The transmitter has a stage in which modulation occurs -- the receiver has a stage in which demodulation occurs. Demodulation is sometimes called detection; it is the recovery of the audio signal from the modulated signal. You will see later that there is a great deal of similarity between the operation of the stage in which modulation occurs and the stage in which demodulation occurs. The transmitter has a microphone which converts the sound into an electrical signal; the receiver has a speaker which converts the electrical signal to a sound signal. Even though the microphone and speaker perform opposite tasks, there is a great deal of similarity between the two.

Now let's look at a simple receiver system. Modern receivers are somewhat more complicated than the one we will describe, but nevertheless millions of receivers like this one have been made.

The Antenna. The radio receiving antenna is a much simpler device than the transmitting antenna. The signals transmitted by modern broadcast transmitters are so strong that only simple receiving antennas are needed. A simple outside antenna may be made from a wire 25 to 50 feet long mounted between two poles. Most modern radio receivers do not need any outside antenna at all. An indoor antenna made of a coil wound on a powdered iron core is called a ***loopstick***. It is mounted in the rear of the receiver, inside of the receiver cabinet. This is all the antenna that is needed to provide satisfactory reception on local broadcast stations. A loopstick mounted in the rear of a receiver is shown in Fig. 22.

The RF Amplifier. The signal picked up by the receiving antenna is quite weak even if the station being received is a fairly strong local station. Before the signal can be used it must be amplified.

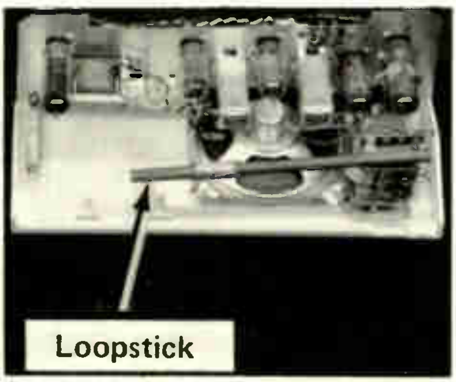


Fig. 22. Loopstick mounted in the rear of a receiver.

This amplification is usually carried out in a stage called an rf amplifier. It is quite similar to the rf amplifier in a transmitter, except it is called an rf voltage amplifier (in a transmitter it is called an rf power amplifier). The rf amplifier in the receiver is designed to increase the strength of the signal voltage picked up by the antenna.

The Demodulator. The amplified signal from the rf amplifier is fed to a stage called the demodulator or detector. This stage separates the audio signal from the rf signal. The rf signal is called the carrier; it carries the audio signal from the transmitter to the receiver. However, it serves little or no useful purpose in the receiver. In the detector stage, the rf signal is separated from the audio signal. The rf signal is discarded so that the signal at the output of the detector is an audio signal. This audio signal is exactly like the audio signal that was originally produced by the microphone in the transmitter.

The Audio Amplifier. The signal at the output of the detector is still a weak signal. Before it can be used to operate a loudspeaker, the strength of this signal must be increased. It is increased by feeding the signal to an audio amplifier, which is similar to the audio amplifier in a radio transmitter. The signal at the output of the audio amplifier is identical to the signal at the input, but much stronger.

The Speaker. A speaker is not very different from a microphone. In fact, sometimes speakers are used as microphones in intercommunications units such as between two offices.

A sketch of a simple speaker is shown in Fig. 23. Notice that the speaker has a magnet like the microphone. Between the poles of the magnet is a coil, and the coil

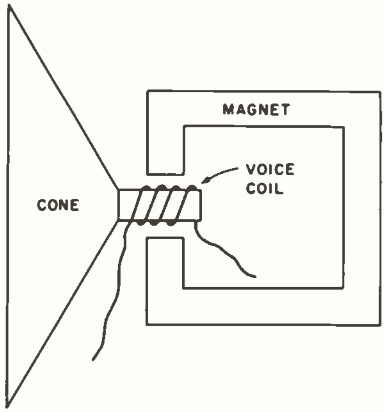


Fig. 23. A simple pm dynamic speaker.

is connected to a diaphragm. However, instead of having a flat diaphragm as in the microphone, the diaphragm is cone-shaped and in a speaker it is called the cone. The cone is fastened to the coil.

When the varying current from the audio amplifier is fed to the coil in the speaker, a varying magnetic field is produced. Depending on the polarity of the field produced by the current flowing in the coil, the field may either aid or oppose the magnetic field produced by the permanent magnet. Since the audio signal is being fed to the speaker, it is actually an ac signal. The polarity of the magnetic field produced by the coil will sometimes aid and sometimes oppose the permanent magnet field. This varying effect will cause the coil to vibrate in and out. Since the coil is fastened to the cone, the cone will vibrate in and out with the coil. The rate of vibration will depend upon the frequency of the audio signal.

The vibrating cone will set the air in front and in back of the cone in motion. The air will vibrate at the same frequency as the cone vibrates. Since the cone is vibrating at the frequency of the original sound signal produced by the microphone, the air around the speaker will be set into vibration at the same frequency. The effect of setting this air into vibration is exactly the same as setting the air into vibration with your vocal chords by speaking. The vibration will be heard as sound, and the sound will be at the same frequency and tone as the original sound that was first uttered in front of the microphone.

The operation of modern speakers is similar to that of the speaker shown in Fig. 23. This type of speaker with a moving coil is called a dynamic speaker. Modern dynamic speakers have a permanent magnet like the one we have shown, and this type is called a permanent magnet dynamic speaker and is usually abbreviated simply a pm speaker. Another type of dynamic speaker uses an electromagnet instead of a permanent magnet. This small coil placed between the poles of the magnet is called the voice coil. Usually the voice coil is wound on a small lightweight form.

The magnets used in pm speakers are strong permanent magnets made of alnico. You will remember that alnico is an alloy of aluminum, nickel, and cobalt which can be used to make extremely strong permanent magnets.

TELEVISION

The transmission and reception of a television signal is not very different from that of a radio signal. However, in television there are two signals to be taken care of, the sound signal and the picture signal. To transmit these two signals through the air, two rf carriers are needed, a picture carrier and a sound carrier. The sound signal in television is called the audio signal as it is in radio and the picture signal is called the video signal.

The sound signal is picked up by a microphone, fed to audio amplifiers and then used to modulate the sound carrier. However, in television, instead of varying the amplitude of the carrier, the frequency of the carrier is varied. This is called frequency modulation and is abbreviated FM. Some radio stations also use this type of modulation. They are called FM stations.

The picture is picked up by a camera that contains a set of lenses similar to the lenses used in a camera that takes photographs. The lenses project the picture on the face of a special tube called a camera pickup tube. This tube has a specially treated face plate on it and the light striking it produces a small voltage. The brighter the light, the more voltage produced. The pickup tube produces a video signal that varies in amplitude with the brightness of the different parts of the picture. The video signal is something like an audio signal; however, it is the electrical equivalent of the picture.

In a color television system, in addition to the signal that contains the brightness information, a color signal is developed, and this signal contains the color information. In other words, the video signal contains the information that tells how bright different parts of the picture are, and the color signal tells what color they are.

The video signal at the output of the pickup tube is fed to a stage called the video amplifier stage where the signal is amplified. In a transmitter, the video signal will be amplified by several video amplifiers. Similarly, in a color TV system, the color signal is amplified by color amplifiers. The color signal is then used to modulate an oscillator which is called a subcarrier oscillator, and this signal is combined with the video signal. The video signal and the color subcarrier signal are then used to amplitude-modulate an rf carrier. Amplitude modulation is used; this is the same type of modulation that is used in the standard radio broadcasting band.

Thus the TV transmitter, instead of transmitting only one rf signal, must actually transmit two rf signals, one to carry the sound signal and the other to carry the video signal along with the color signal if the broadcast is in color. The two signals from the transmitter are fed to a single antenna.

At the receiving installation, the two signals are picked up by one antenna, amplified by rf amplifiers and then separated from the rf carriers by separate detectors, one for the video and the other for the sound. In the case of a color television receiver, the video signal and the color subcarrier are separated from the video carrier and the color signal is then separated from the color subcarrier by a color detector.

The sound signal from the detector is amplified and fed to a speaker. The video signal is amplified and fed to the picture tube. In a black and white TV receiver the picture tube brightness is controlled by the video signal. In a color television receiver the brightness is controlled by the video signal which is mixed with the color signal to produce the original colors picked up by the camera.

This is only a brief run through of a

television transmitter and receiver. There are many details that have been simplified, many more that have been omitted, but you need not be concerned about them at this time. Keep in mind that the operation of the video portion of a TV transmitter is not very different from that of the audio portion of the transmitter. There are many differences in details, but the basic principles are the same.

SUMMARY

This brief description of a broadcasting system gives you a birdseye picture of what happens in a radio transmitter and a radio receiver. We do not expect you to remember the details at this time, but having a general idea of what takes place will help you with later lessons. Remember that an rf stage amplifies a radio frequency signal and that an audio stage amplifies an audio signal. Remember what an rf signal is, what an audio signal is and what a video signal is. You should also remember that the stage that separates the audio signal from the rf carrier and the stage that separates the video signal from the rf carrier are both called detector stages. You should remember the basic principles of the operation of a microphone and a speaker and also the schematic symbols for a microphone and an antenna.

SELF-TEST QUESTIONS

(v) How many kilohertz are there in a megahertz?

(w) Write 1900 kHz in megahertz.

(x) What is the name given to the electrical signal produced by sound striking a microphone?

(y) What is an audio stage?

(z) What is the name given to the stage in a transmitter that actually generates the radio frequency signal?

(aa) What is the name of the device used to feed an rf signal from the transmitter to the antenna?

(ab) What is the name given to the stage that separates the audio signal from the rf carrier?

(ac) What is the name of the small coil placed between the poles of the magnet in a speaker?

(ad) What are the names of the two signals that must be transmitted in a black and white television system?

(ae) What type of modulation is used in the sound section of a TV transmitter?

(af) What type of modulation is used in the video section of a TV transmitter?

Electronics in Industry

We have already mentioned that there are many uses for electronics in industry and that the number of applications is growing daily. Electronics is used in oil refineries to control the various steps in the refinement of crude oil. It is used in the livestock feed industry to control the mixing of grains and the preparation of feed for farm animals. It is used in the factories to control the operation of precision machines, to inspect the finished product coming off the assembly lines and to count the output of high-speed automatic machines. It is used by railroads to automatically guide loaded and/or empty cars in switching operations and to control the speed of the car so that it hits the cars already standing on the track at just the right speed to couple to the car without damaging the cars or their contents. We have all been thrilled in the past few years by the amazing feats performed by space ships sent to the moon to take television pictures of the moon, by the launching and orbiting of the ships and recovery of the astronauts. All these phenomenal feats have been possible due to many electronic devices at the control stations and aboard the space ships. In spite of all the advances we have made, the chances are that in the next ten years we will see even greater advances and even more opportunities in the field of electronics.

CHANGING AC TO DC

One of the important uses of electronics in industry is converting ac to dc. This can be accomplished by large electronic tubes designed for this type of service and by solid-state devices.

It is much more convenient to generate and transmit ac than it is dc. The reason for this is that ac can be transmitted at very high voltages. The higher the voltage is for a given amount of power to be transmitted, the lower the losses in the transmission lines will be. AC voltages can be conveniently increased or decreased by means of a transformer, whereas dc voltages cannot be changed from one value to another conveniently.

The Transformer. You have already briefly studied a simple transformer. A transformer consists of two coils wound on a common core. In the preceding lesson you learned that if a battery was connected to one winding and a switch inserted into the circuit and the switch opened and closed rapidly, the changing magnetic field set up in the one winding would induce a voltage in the second winding.

We call the winding to which the battery is connected the primary winding and we call the other winding the secondary winding. These names are easy to remember; remember that primary is first, and secondary is second.

If instead of a battery and switch, we connected ac voltage to the primary winding on the transformer, the ac voltage will cause an alternating current to flow in the primary winding. When the alternating current flows in one direction it will build up a magnetic field and the changing lines of force, as the field is built up, will cut the turns of the secondary winding and induce a voltage in these turns. As the ac current collapses and then begins to flow in the opposite direction, the magnetic field being produced will be continually changing. As the field builds up in the opposite direction, a voltage of the opposite polarity will be induced in the secondary winding.

The voltage that will be induced in the secondary winding will depend upon how many turns there are on the secondary winding. If the secondary winding has the same number of turns as the primary winding, the voltage induced in the secondary will be equal to the voltage applied to the primary. If the secondary winding has twice as many turns as the primary winding, the voltage induced in the secondary will be twice the voltage induced in the primary. A transformer of this type is called a step-up transformer because it steps the voltage up to a higher value. On the other hand, if the secondary winding has only half as many turns as the primary, the voltage induced in the secondary will be only half the voltage applied to the primary. This type of voltage is called a step-down transformer because it steps down the voltage.

Thus by means of a step-up transformer an ac voltage generated by a generator can be stepped up to a very high value. It can then be transmitted to a distant point by means of high-voltage transmission lines and at that point stepped down by means of a step-down transformer. This is why it is more convenient to transmit alternating current than direct current.

Alternating current is changed to direct current by means of rectifiers. A rectifier is a device that will let current flow through it in one direction, but will not let it flow through it in the opposite direction. You will learn about tube rectifiers now and about other types later.

Half-Wave Rectifiers. A half-wave rectifier is a device that will allow current to flow during only one half of each cycle. Remember that the ac cycle we looked at before in Fig. 13 had a positive half and a negative half. A half-wave rectifier can be connected to allow current to flow either during the positive half or during the negative half of the cycle, but it will not let current flow during both halves of the cycle.

A schematic diagram of a half-wave rectifier circuit is shown in Fig. 24. You are already familiar with the schematic symbols used on this diagram. T_1 and T_2 are iron-core transformers, and V_1 is a tube with a filament and a plate. Notice that the tube is drawn upside down. You will find them drawn in any position on schematic diagrams, but the pointed symbol is always the filament and the straight line is the plate. On the schematic diagrams the dots where several connecting lines meet indicate a connection. The crossovers without a dot indicate that there is no connection.

Now let us study Fig. 24. The ac power from the power line is fed to the primary

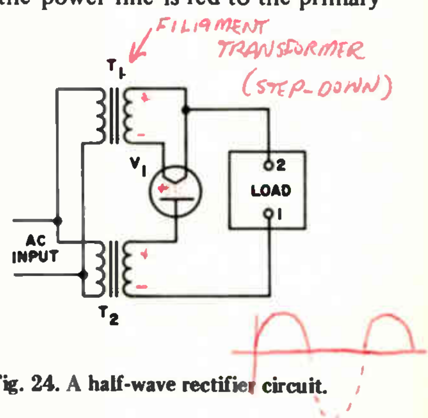


Fig. 24. A half-wave rectifier circuit.

winding of the transformers marked T_1 and T_2. T_1 is a step-down transformer. The winding on this transformer provides the voltage necessary to heat the filament of the diode tube. This transformer is often called a filament transformer because it supplies the power required to heat the tube and serves no other useful purpose in the circuit.

Transformer T_2 may be either a step-up or step-down transformer depending upon the dc voltage required by the load. If the dc voltage needed is higher than the power line voltage, a step-up transformer is used, whereas if it is lower, a step-down transformer is used. The block marked "load" on the diagram represents whatever is going to use the dc power produced. This might be a number of storage batteries we are charging, or it could be a bath which is being used to refine copper or aluminum in a refinery. You will see the word load used frequently in this way.

Now let us see how the half-wave rectifier works. Refer to Fig. 25 as you read the explanation. We have simplified the figure by leaving out the filament transformer T_1 and the primary winding

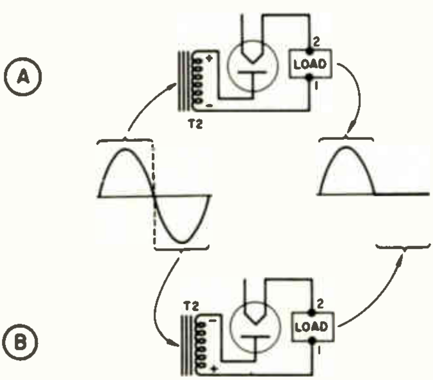


Fig. 25. How a half-wave rectifier works.

of T_2. The filament transformer T_1 heats the tube filament, but does not enter into the operation of the rectifier in any other way.

Looking at Fig. 25A we see the first half cycle which is the positive half cycle. At the left of the drawing we see the ac voltage across the secondary of T_2. In the center we see the polarity of the voltage across the secondary of T_2 and at the right we see the voltage that will appear across the load. The voltage will cause a current flow through the load. During this half cycle, the end of the secondary winding of T_2 which is connected to the plate of the diode tube is positive and the other end is negative. When this happens, current will flow from the lower end of the transformer winding to terminal 1 of the load, through the load to terminal 2 and to the filament of the tube. You know that the red-hot filament will emit or give off electrons. These electrons will be attracted to the plate of the tube by the positive voltage on the plate. Therefore, the electrons will flow from the filament to the plate of the tube to the positive end of the transformer. Since this is a complete circuit, current can flow.

However, when the polarity of the voltage across the secondary of T_2 reverses, we will have the negative half cycle shown in Fig. 25B. The polarity of the secondary of T_2 is shown. Notice that the end of the secondary winding connected to the plate of the tube will be negative and the other end will be positive. Now let us see how current will have to flow during this part of the cycle.

You will remember that current flows from the negative terminal of the voltage source through the external circuit and back to the positive terminal of the source. This means the electrons will have to flow from the negative end of the secondary winding of the transformer to the plate of the tube. They could do this, but then they would have to flow from the plate of the tube to the filament. However, there is no way the plate can give off electrons. Furthermore, electrons will not flow from the filament of the tube to the plate because they will be repelled by the negative voltage on the plate. The plate will be negative because the side of the transformer it is connected to is negative during this half cycle. Since electrons cannot get across the tube there is no complete circuit and, therefore, there will be no current in the circuit.

The current in a half-wave rectifier of this type will look like the drawing in Fig. 25. Notice that during the first half of the cycle when the plate is positive there will be current through the circuit. During the next half cycle when the plate is negative there is no current. This chain of events will continue so that during the third half cycle when the plate becomes positive again, current will flow and then during the fourth half cycle again there will be no current. The current will flow in pulses, with one pulse for each cycle. Again, this type of dc is called a pulsating dc and the rectifier is called a half-wave rectifier because it rectifies only one half of each cycle.

In a practical circuit the need for two transformers can be avoided by putting two secondary windings on one transformer. One winding as a step-down winding provides the voltage needed to operate the filament of the rectifier tube. The other winding may be either a step-up or a step-down winding depending upon the dc voltage required by the load.

SILICON RECTIFIERS

In many applications silicon rectifiers have replaced diode tubes. The advantage of a silicon rectifier is that it does not have a filament and hence does not use any power to heat the filament such as a tube does. In addition, silicon rectifiers are quite small and can pass rather large currents for their size. They have an additional advantage in that there is very little voltage lost across the rectifier, whereas in a tube there will be some voltage lost across the tube. A silicon rectifier is made of two different types of silicon placed together to form a junction. This junction will permit current to cross it in one direction, but will not permit it to cross in the other. In the forward direction there is practically no resistance in the junction and therefore the current flows freely in that direction. In the reverse direction, the junction offers such a very high resistance that practically no current crosses it.

A schematic diagram of a half-wave rectifier using a silicon rectifier is shown in Fig. 26. The arrows indicate the

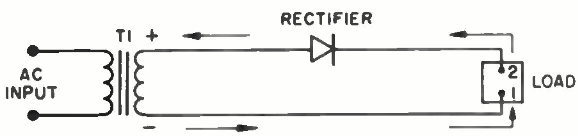


Fig. 26. A half-wave rectifier using a silicon rectifier.

direction in which current will flow during the half cycle the rectifier passes current.

Notice that the arrows indicating the direction of current flow point in the opposite direction to the arrow used as part of the symbol in the silicon rectifier. The reason for this is that the symbol dates back to the early days of electricity when engineers and scientists thought that current flowed from the positive terminal of a battery or generator through the load and back to the negative terminal. Hence the symbol was drawn with the arrow indicating this direction of current. Now we know that current flows in the opposite direction, but the symbol has been carried over to indicate a solid-state rectifier such as a silicon rectifier and the arrow points in the wrong direction. Remember this symbol, it is used to represent all types of solid-state devices that will permit current flow in only one direction. It is used to represent detectors made out of germanium which can be used in radio and television receivers; it is used to represent selenium rectifiers which are another type of solid-state rectifier; and it is used to represent copper-oxide rectifiers which are used in meters. You will learn more about these devices later. They are often called diodes because there are two types of material used in them.

SUMMARY

At this time you need not remember all the details of how a rectifier operates. However, you should remember that a rectifier is a device that will permit current to flow through it in only one direction. By using a rectifier, alternating current can be changed to pulsating direct current. Remember that a half-wave rectifier rectifies only half of each cycle so that you'll get a pulse during one half cycle and no current during the next half cycle.

SELF-TEST QUESTIONS

(ag) What is the name given to a transformer where the secondary voltage is higher than the primary voltage?

(ah) If a transformer has fewer turns on the secondary winding than it has on the primary winding, will the secondary voltage be equal to, greater than or less than the primary voltage? What is the name given to this kind of transformer?

(ai) In a half-wave rectifier circuit, in which direction will current flow through a diode tube?

(aj) What must the polarity of the voltage on the plate of a diode rectifier tube be in order for current to flow in a half-wave rectifier circuit?

(ak) Draw the schematic symbol for a silicon rectifier and by means of an arrow above it indicate in which direction the current will flow through the rectifier.

(al) When two connecting lines cross and there is a dot placed on the junction, what does this indicate?

Electronics for Computers

Almost all of the electronic circuits used in communications equipment, industrial applications and in computers operate from direct current. While this dc could be obtained from a set of batteries, most often it is obtained from a power supply that consists of a power transformer and rectifier similar to that discussed in the previous section. The transformer takes the ac line voltage and steps it up or down to the required voltage level. The rectifier then converts this ac into a pulsating dc. This pulsating dc is then fed through a filter circuit to smooth it into a clean pure dc signal very similar to that obtained from a battery. The resulting dc voltage is then used to power the various electronic circuits used in communications, industrial and computer equipment.

The sine wave is the basic electrical signal used in communications and industrial electronic equipment. In communications equipment, oscillator circuits are used to generate a sine wave signal and produce a carrier signal that is radiated by the transmitter and detected by the receiver. In computer circuits, however, the sine wave is not the basic electrical signal. Computers use another class of signals known as pulse signals. Let's take a look at some of these interesting signals and see how they are produced and used in computers.

PULSE SIGNALS

A pulse signal is an electrical voltage or current that switches rapidly from one voltage level to another and back again. A

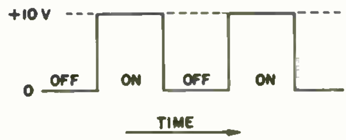


Fig. 27. A simple pulse signal.

simple pulse signal is shown in Fig. 27. Notice that this signal switches rapidly from a zero volt level to a 10 volt level and then back again to zero, repeating itself periodically. The resulting wave shape of this signal is approximately square and, therefore, this pulse signal is often referred to as a square wave or more generally as a rectangular wave. This signal is basically a pulsating dc signal similar to the output of the half-wave rectifier you studied in the previous section. The difference between this signal and the rectifier output is the wave shape.

A pulse signal like this is relatively easy to generate. In fact, we could generate it very simply with the battery and switch circuit shown in Fig. 28. When the switch is open, or off, no voltage appears between points A and B. However, when we close the switch or turn on the circuit, the battery voltage appears between points A and B. If we turn the switch off

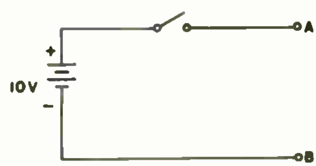


Fig. 28. One means of generating a square wave.

and on at a rapid rate, it will produce a square wave similar to that shown in Fig. 27.

While pulse signals are sometimes produced with a circuit as simple as that shown in Fig. 28, they are usually produced by electronic circuits known as multivibrators. A multivibrator generates a very fast high frequency square wave whose frequency, amplitude and pulse width can be varied over wide ranges.

Any pulse signal is usually specified by these three characteristics: frequency, amplitude and pulse width. Frequency is the rate at which the pulses occur. The amplitude of the pulse is the amount of voltage or current that it represents when it is "on". Pulse width designates the amount of time that the pulse is "on".

There are a wide variety of different pulse signals that you may encounter in computers and in other electronic equipment. The signal in Fig. 29A is a rectangular wave that switches from 0 to +5 volts. This signal is similar to that shown in Fig. 27 with the exception that the amplitude of the voltage is less and the off and on times are unequal. It is quite common to have unequal off and on periods in a pulse signal. Here we show the on period being longer than the off period. However, just the opposite could be true in another waveform.

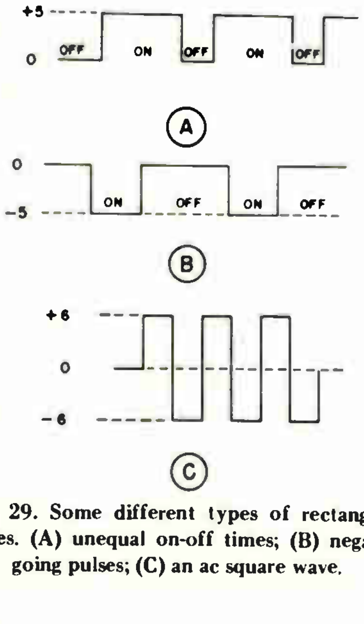


Fig. 29. Some different types of rectangular pulses. (A) unequal on-off times; (B) negative going pulses; (C) an ac square wave.

Fig. 29B shows another rectangular waveform where the voltage switches from 0 volts to –5 volts. It is easy to generate either positive or negative voltage signals with pulse circuits.

In Fig. 29C we show yet another type of pulse circuit that switches between two voltage levels, +6 volts and –6 volts. This type of pulse signal is really an ac signal since it produces an alternating current. When the pulse is positive, the voltage will cause electrons through a circuit to flow in one direction. When the voltage is negative, electrons will flow through the circuit in the opposite direction.

The pulse waveforms we have shown here are typical of what you might encounter in computer equipment but other variations are possible. Pulse signals like these are also used in television, telemetry and communications equipment. Just keep in mind the basic fact that a pulse signal usually switches rapidly between two voltage or current levels rather than varying smoothly and continuously like a sine wave does.

LOGIC CIRCUITS

There is a special class of electronic circuits known as logic circuits that are used to process or manipulate the pulse signals you have just studied. Such cir-

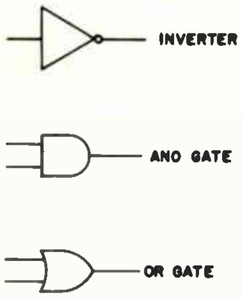


Fig. 30. The three basic digital logic circuits.

cuits are also referred to as digital circuits. The three basic logic circuits are the inverter, the AND gate, and the OR gate. The symbols representing these three basic digital circuits are shown in Fig. 30. Let's see how these circuits work. All of these circuits operate on pulse signals similar to those shown in Fig. 27 and 29.

An inverter circuit does exactly what it says, it inverts a pulse signal so that the inputs and outputs are going in opposite directions. We say that the signals are complements of one another. For example, if a pulse switching from +5 volts to 0 volts is applied to the input of an inverter as shown in Fig. 31A, the resulting output will be a signal that switches

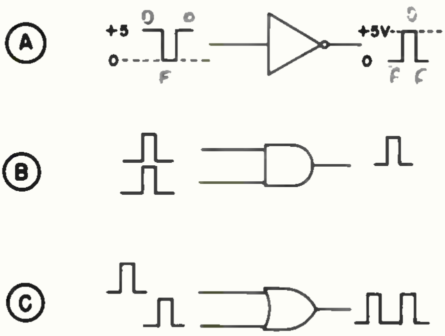


Fig. 31. (A) inverter operation; (B) AND gate operation; (C) OR gate operation.

from 0 to +5 volts at the output. A positive-going pulse at the input to an inverter produces the negative-going pulse on the output.

An AND gate is a logic circuit that produces an output pulse if two or more pulses are applied to its inputs simultaneously. See Fig. 31B. An AND gate can have two or more inputs and a single output. The gate does not produce an output signal unless input signals are applied to all of its inputs simultaneously. If an input signal occurs on only one of the input lines, no output pulse will be produced. The AND gate is commonly referred to as a coincidence gate since it produces an output only when the two inputs are coincident in time.

Fig. 31C shows how the OR gate works. The OR gate also has one output and can have two or more inputs. It produces an output any time a pulse occurs on any of its inputs. In Fig. 31C we show two pulses occurring at different times at the two inputs. An output pulse is produced for each of the input pulses.

Any digital logic function can be performed by just these simple circuits. More complex operations than those described here are performed by combining these three circuits. Any digital computer is made up of nothing but a large quantity of these three basic types of circuits. A digital computer or other large digital system may be very complex. However, if you understand the operation of the three simple circuits discussed here, you can easily learn how the more complex systems work. This is true, of course, of any piece of electronic equipment. If you learn how the basic simple circuits such as oscillators, amplifiers and power supplies work, you will then be able to understand more complex pieces of equipment. In your NRI lessons to come, you will study these important fundamental circuits to prepare you for more complex equipment.

Now answer the self-test questions on this section and then do the lesson questions. Remember that if you have difficulty with the self-test questions it is an indication you need to go over the section again. Do not hesitate to spend any time you need to review. Learning the basic fundamentals you are studying now is the most important part of your course. If you understand these fundamentals, you will be able to build on them in later lessons and you will find that as you go on your course becomes easier and easier. On the other hand, if you do not master the fundamentals covered in the early lessons you will find it difficult to grasp some of the ideas presented in later lessons.

SELF-TEST QUESTIONS

(am) What is the main difference between a pulse signal and a sine wave signal?

(an) Name three characteristics by which a rectangular pulse signal is usually specified.

(ao) Which type of logic circuit produces an output pulse if input pulses occur on either of its two inputs?

(ap) True or False? A pulse signal can be either ac or pulsating dc.

Answers to Self-Test Questions

(a) 1-1/2 volts.

(b) 22-1/2 volts. Each cell has a voltage of 1-1/2 volts and since there are fifteen cells, the voltage will be 1-1/2 × 15 = 22-1/2 volts.

(c) 1-1/2 volts. When cells are connected in parallel the voltage of all the cells must be equal and the total voltage of the cell will be equal to the voltage of one cell. Connecting cells in parallel does not increase the voltage -- it increases the current capabilities of the battery.

(d) +4-1/2 volts. The battery will be marked essentially the same as the battery in Fig. 4. Since one terminal is marked with a – sign, we consider this terminal as the reference point. The center tap has three cells between it and the reference point and therefore the voltage will be 4-1/2 volts. It will be marked + to indicate that this terminal is positive with respect to the negative terminal.

(e) A longer shelf life and a higher current capacity.

(f) 1.35 volts or 1.4 volts, depending upon the materials used in the mercury cell.

(g) 1.5 volts. The output voltage of a manganese cell is approximately the same as the output voltage of a dry cell.

(h) The manganese cell has not replaced the dry cell because it is more expensive than the dry cell, so in many cases the economy of the dry cell overrules the advantages of the manganese cell.

(i) The lead-acid cell and the nickel-cadmium cell.

(j) The nickel-cadmium cell provides longer life and more maintenance-free performance than the lead-acid cell because it can be sealed and hence water does not escape from it when it is charged.

(k) Direct current and alternating current.

(l) DC voltage.

(m) AC voltage.

(n) The voltage wave supplied by the power company is called a sine wave.

(o) The waveform shown in Fig. 13B represents one cycle. In a cycle of alternating current, the voltage between the two terminals of the generator starts at zero and then builds up to a maximum value with one polarity and then drops back to zero; it then builds up to a maximum value with the other polarity and then drops back to zero. During the next cycle it will simply repeat the first cycle.

(p) 60 cycles per second means that we have 60 complete cycles in a second. This means that the voltage starts at zero, builds up to a maximum with one polarity, drops back to zero, builds up to a maximum value at the opposite polarity and drops back to zero a total of sixty times in a second.

(q) A commutator.

(r) A simple dc generator such as shown in Fig. 14 would produce a pulsating dc such as shown by the waveform in Fig. 16. A pulsating dc of this type is normally not desired and the excessive pulsating can be eliminated by means of an armature such as shown in Fig. 17. In this armature a number of series-connected coils are used and a more even dc is produced.

(s) When we say that the ac current is 1 amp we mean that the ac current is the equivalent of 1 amp of direct current. In other words, if an ac current of 1 amp is flowing through the coils of a heater, it will produce the same amount of heat as a dc current of 1 amp flowing through the same coils would produce.

(t) When we give an ac voltage and do not specifically mention that it is the peak value, we always assume that the value is meant to be the effective or rms value of the voltage. If we want to give the peak value of an ac voltage, we should specifically state that it is the peak value, otherwise it would be assumed to be the effective or rms value.

(u) 280 volts. The peak value of an ac voltage is 1.4 times the rms value. Thus if the rms value is 200 volts, the peak value will be: 200 X 1.4 = 280 volts.

(v) There are 1,000 kilohertz in one megahertz.

(w) 1900 kHz = 1.9 MHz.

(x) An audio signal.

(y) An audio stage is a stage designed to amplify an audio signal. The stage may consist of a tube and a number of components or it may consist of a transistor and a number of components.

(z) The oscillator stage.

(aa) The transmission line.

(ab) A detector stage or a demodulator stage. Both names are used.

(ac) A voice coil.

(ad) The audio signal which carries the sound information and the video signal which carries the picture information.

(ae) Frequency modulation (FM) is used in the sound section of a TV transmitter.

(af) Amplitude modulation (AM) is used in the video section of a TV transmitter.

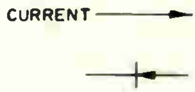
(ag) A step-up transformer.

(ah) Less than; a step-down transformer.

(ai) Current flows from the filament or cathode to the plate.

(aj) The plate must be positive.

(ak)



(al) It indicates a connection between the two circuits.

(am) A pulse signal switches abruptly from one voltage level to another while a sine wave varies smoothly or continuously between its two peak levels.

(an) Frequency, amplitude and pulse width.

(ao) An OR gate.

(ap) True.

YOU'VE GOT COURAGE

An eight year old boy was discussing arithmetic with a friend. Said Johnny: "Teacher's gonna start us on subtraction tomorrow; wish I could stay home." But Harry laughed outright at Johnny's fears, and said; "Aw, I've had that and it's easy once you get started; what I'm worrying about is multiplication!"

It is natural even for grown men to feel like these boys – to fear most the things about which they know the least.

Did you know that some of the world's best speakers are always afraid when they get up to make a speech before a strange audience? Their courage gets them started, and in no time at all their fear changes to a confident enthusiasm which makes their talk a big success.

You've got courage! Use your courage to overcome normal fears, to carry you into each new subject and carry you over each difficulty. In no time at all, you will be looking forward to new subjects with intense interest – you will be eager to tackle new problems. Remember that each conquered difficulty brings you one step closer to your goal of SUCCESS IN ELECTRONICS!

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-202

ACHIEVEMENT THROUGH ELECTRONICS

nri

CURRENT, VOLTAGE, AND RESISTANCE

B103

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

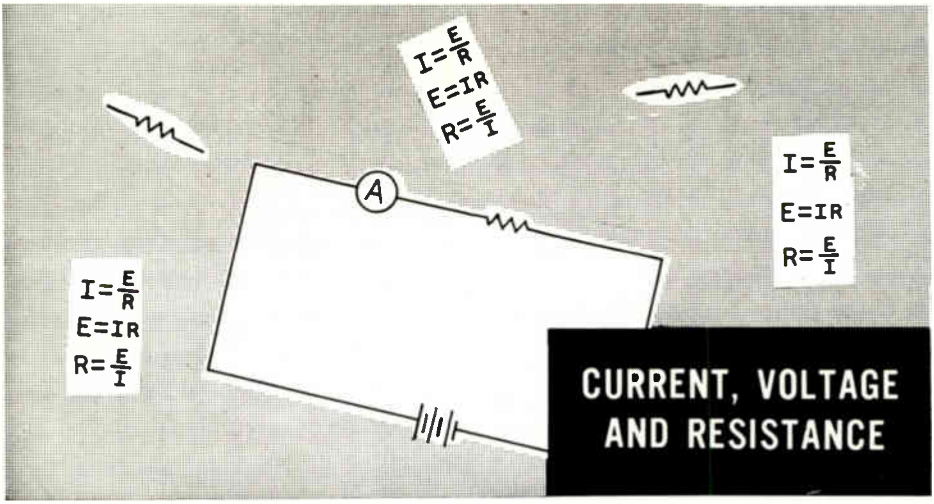
CURRENT, VOLTAGE AND RESISTANCE

B103

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

PII3M1174 1974 EDITION Litho in U.S.A.



CURRENT, VOLTAGE AND RESISTANCE

Have you ever looked into the back of a color television receiver? If you have, you probably wondered how it would be possible to identify the parts in the set and to trace out the circuits. You may have wondered how an experienced technician can find a defective part among the maze of parts and circuits in the receiver. The technique that he uses is the same as the technique that you are going to learn now. Most of the circuits in a color television receiver or in any other piece of electronic equipment are basically rather simple circuits. The difficulty comes from the fact that so many of them are used together. However, once you learn how to trace simple circuits and understand what is happening in the circuit you will find that you will be able to apply that knowledge to more complex circuits, and also to larger groups of simple circuits when they are used together.

Your first two lessons introduced you to a number of important things in the field of electronics. You learned the important laws of electric and magnetic charges: like charges repel, unlike charges attract. In magnets like fields repel and unlike fields attract. You also learned that an electric current was a movement of electrons. You already know that an electron has a negative charge and that it will be attracted by a positive charge and repel other negative charges.

You learned about direct and alternating currents and studied batteries and generators. You also took a quick look at a broadcast system and learned what is meant by an audio signal and a radio frequency signal.

We will go into more detail later on the various things covered in your first two lessons. These lessons were primarily introductory lessons to get you started. In this lesson we will start going into considerable detail. You will learn that the units of electrical measurement - the volt, the ampere, and the ohm - are in some cases too small and in other cases too large. You'll learn how to convert them into practical sizes when necessary.

You will study simple circuits and learn how voltage, current and resistance are related in these circuits. You will study Ohm's Law; this is a very simple law, but it is probably the most important rule or law in electronics. You will use Ohm's Law over and over again as long as you are in the electronics field.

As you study this lesson, keep in mind that it is perhaps the most important lesson in your entire course. You must understand basic simple circuits thoroughly. If there is anything you do not understand, be sure to go over that section of the lesson several times. If you can't work it out for yourself be sure to write to NRI and ask for help. Our instructors will be glad to help you; they all agree that this is a very important lesson and will do everything they can to be sure that you understand the entire lesson.

Pay particular attention to the self-test questions. These questions will tell you whether or not you have learned all you should have about each section of the lesson as you go along. If you find there is a self-test question you cannot answer, this means you need to restudy that section of the lesson. Don't be afraid to go back and spend whatever extra time may be necessary. A little extra time on this lesson will pay great dividends in later lessons not only in helping you understand these lessons better, but it will also make it easier for you and save time in the long run.

Now let us go ahead and learn more about current, voltage and resistance. We will study current first, then voltage, and finally resistance. In each case, we will review the important points that you already know, and then go ahead and expand that knowledge.

Current

You have already learned that an electron is part of an atom. An electron has a negative electrical charge. All electrons have exactly the same negative electric charge. Electrons repel each other because like charges repel. Electrons are repelled by all negative charges and attracted by all positive charges.

An electric current is the movement of electrons in the circuit. There are two kinds of electric current, direct current and alternating current. Let us discuss direct current first.

DIRECT CURRENT

When a battery is connected to an electric circuit, electrons are repelled from the negative terminal of the battery and attracted by the positive terminal of the battery. Electrons move in one direction through the circuit. This type of current flow is called direct current. We usually refer to it simply as dc.

In a simple circuit such as shown in Fig. 1, it is important for you to remember that the current flow is the same in all parts of the circuit at all times. The number of electrons leaving the negative terminal

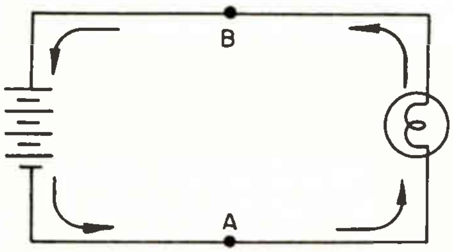


Fig. 1. A simple dc circuit.

of the battery is exactly the same as the number of electrons passing point A on the wire connecting the battery to the lamp. The number of electrons flowing through the lamp is the same as the number of electrons passing point B on the other wire, and the same number of electrons are moving into the positive terminal of the battery.

Another important point that you must remember about an electric current is that the instant the circuit is closed, the electrons start moving in all parts of the circuit. One electron does NOT leave the negative terminal of the battery, hit another, and so forth, and thus start movement in the circuit after a certain period of time; each electron starts movement immediately the moment the circuit is completed.

The unit of current flow is the ampere which we usually abbreviate amp. A current of 1 ampere represents a certain number of electrons moving past a point in the circuit in a second. If twice as many electrons are moving past that point in the circuit, the current flow is 2 amperes, and if ten times the number of electrons are moving past the point in the circuit, the current flow is 10 amperes.

Direct current is widely used in electronic equipment. Current will flow through a vacuum tube or a transistor in only one direction and therefore the current used to operate these devices is direct current. There are other applications of direct current in industry. Direct current is used in purifying metals, in plating operations and in running some motors that require precise controls. Direct current motors can be more closely controlled than alternating current motors.

ALTERNATING CURRENT

The electric current supplied to homes for lighting, for cooking and operation of small appliances is alternating current. An alternating current flows first in one direction and then in the other direction. The current starts at zero amplitude and builds up to a maximum value which we call the peak value, then drops back to zero, then builds up to a peak value flowing in the opposite direction, and then drops to zero again. This action of flowing first in one direction and dropping back to zero value, then building up to a maximum value in the opposite direction and then once again dropping back to zero is called a cycle. The current supplied by most power companies is 60-cycle current. This means that it goes through 60 cycles per second.

Alternating current is also measured in amperes. However, since the number of electrons flowing past a point in the circuit is continually changing when we have alternating current, we use a somewhat different method of expressing the value of the current. The alternating current is compared to direct current. If the alternating current flowing through a heat-producing device produces the same amount of heat as a direct current of 1 ampere produces, we say that the ac current is 1 ampere. We call this the effective value or the rms value.

The change in amplitude of an alternating current follows a pattern or waveshape which is known as a sine wave. A typical sine wave showing one complete cycle is shown in Fig. 2. In this figure the rms value of the alternating current is 1 amp. As you can see, for part of the cycle the actual current flowing will be less than 1 amp and during part of the cycle it will be greater than 1 amp. The actual peak or highest value that the current reaches is 1.4 × the rms value. In the example shown, where the rms value is 1 amp, the peak value will be 1.4 amps.

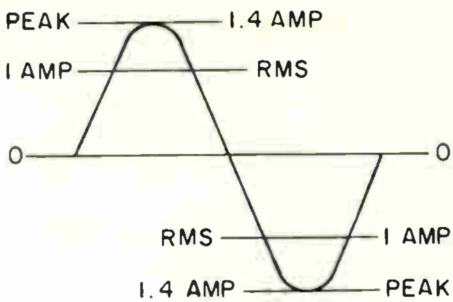


Fig. 2. A sine wave showing rms and peak values.

In the first half cycle, we have drawn the waveshape above the horizontal line representing zero current. This half of the cycle is referred to as the positive half cycle. The other half cycle is referred to as the negative half cycle. Notice, the negative half is exactly the same as the positive half. The current reaches the same rms and the same peak values. We simply call one a positive half cycle and the other a negative half cycle to indicate that the current is flowing in opposite directions during the two half cycles.

To convert rms values of current to peak values you multiply the rms value by 1.4. If you have the peak value of the current and want to convert it to the rms value, you simply divide by 1.4.

THE MILLIAMPERE

While the unit of current measurement is the ampere, in electronics this unit is often so large that it is cumbersome. For example, in an audio amplifier stage using either a tube or a transistor, the current flow may be only a few thousandths of an ampere. In a case where the current flow was three thousandths of an ampere we could write this as 3/1000 or we could write this value as a decimal, in which case it would be .003.

However, rather than do this it is much more convenient to express the current flow in milliamperes. A milliampere is one thousandth of an ampere. Thus a current flow of three thousandths of an ampere will be 3 milliamperes.

There are one thousand milliamperes in an ampere. Therefore to convert amperes to milliamperes you simply multiply by one thousand. You do this by moving the decimal point three places to the right. For example, suppose you have a current of 5 amperes. We would normally simply write this as 5 amps. However, we can also write it as 5. amps. Now to move the decimal point three places to the right we simply add three zeros to the right of the decimal point, and then move the decimal point to the right of the three zeros. Thus 5.000 amps becomes 5,000. milliamps.

We seldom use milliamperes to express currents that are over 1 ampere, but when the current is some fraction of an ampere the milliampere becomes particularly useful. For example, suppose the current flow in a circuit is .05 amperes. To change this to milliamperes we move the decimal point three places to the right. To do this, write .05 amps as .050 amps and then move the decimal point three places to the right. Thus .05 amps = .050 amps = 50 milliamperes.

If you have a current in milliamperes and want to convert it to amperes you move the decimal point three places to the left, adding zeros as necessary. For example, to convert 47 milliamperes to amps we first write 47 milliamperes as 47. milliamperes. Now we move the decimal point three places to the left by moving it past the 7 and past the 4. Now we add a 0 so we can move the decimal point three places and get .047 amps. Another example, suppose the current is 7 milliamperes and we want to convert this to amperes. We write 7 milliamperes as 7. milliamperes. Next, we move the decimal point three places to the left; we move it past the 7 and then add two zeros to the left of the 7 and get .007 amps.

You will deal with milliamperes a great deal in electronics, and often you will have to convert them to amperes. Remember the rules for converting back and forth. To convert milliamperes to amperes you move the decimal point three places to the left. To convert amperes to milliamperes you move the decimal point three places to the right. In effect, when you convert from milliamperes to amperes by moving the decimal point three places to the left, you are dividing by 1000. When you convert amperes to milliamperes by moving the decimal point three places to the right, you are multiplying by 1000.

Since we use milliamperes so often in electronics it is convenient to have an abbreviation for this rather long word. We often abbreviate milliampere milliamp and to make it plural we simply add an s. An even more convenient abbreviation is ma. Thus 27 milliamperes can be abbreviated 27 milliamps or 27 ma.

You might think that converting from amperes to milliamperes and from milliamperes back to amperes is difficult and something that you are not used to doing. This is not true. Whether you realize it or not you are doing conversions of this type all the time. For example, suppose somebody gave you four hundred cents. It is not too likely that you would say that you had four hundred cents. Chances are that you would convert cents to dollars by moving the decimal point two places to the left and say you had $4.00. However, if you did have $4.00 and wanted to change it to cents you would move the decimal point two places to the right and know that you should have four hundred cents. The conversion back and forth between amps and milliamps is exactly the same except that we have one additional place. To convert from the larger unit, dollars in one case and amperes in the other, to the smaller unit, cents in one case and milliamperes in the other, we move the decimal point to the right. To convert from the smaller unit (cents in the one case and milliamperes in the other) to the larger unit (dollars in the one case and amps in the other case) we move the decimal point to the left. Remember the dollars and cents conversion and remember that you have one extra place and you will have no difficulty changing back and forth between amps and milliamps; it is that easy.

THE MICROAMPERE

In some circuits, even the milliampere is too large a unit to conveniently express current flow. Thus we have an even smaller unit, the microampere, which we abbreviate microamp. The microamp is one millionth of an ampere. It is one thousandth of a milliamp.

To convert amps to microamps you move the decimal point six places to the right. Remember it is the same as converting amps to milliamps except that you move the decimal point six places instead of three places. To convert microamps to amps you move the decimal point six places to the left. This is the same as converting milliamps to amps, or cents to dollars except that you move the decimal point six places to the left.

Even the abbreviation microamp is somewhat long and inconvenient, and since we have already used the letter ma for milliamps, we use the Greek μ which looks something like our u to abbreviate microamps. We write it μa.

Sometimes you will want to convert microamperes to milliamperes and vice versa. To do this, you move the decimal point three places. To go from the larger unit, milliamperes, to the smaller unit, microamperes, move the decimal point to the right, and to go from the smaller unit to the larger unit move it to the left.

In Fig. 3 we have shown a number of examples of conversions from one unit to another. Before going ahead study this figure and the conversions and then try to do them yourself. This is the best way to learn how to convert from one unit to another. Once you learn how to do this you

LARGE TO SMALL	SMALL TO LARGE
Dollars To Cents $1 = 100 cents We moved the decimal point 2 places to the right. $1 = $1.00 $1.00 = 100 cents	**Cents To Dollars** 1000 cents = $10.00 = $10. We moved the decimal point 2 places to the left. 1000. ¢ = $10.00
Amps To Milliamps 1 amp = 1000 milliamps We moved the decimal point 3 places to the right. 1 amp = 1.000 amp 1.000 amp = 1000 milliamps	**Milliamps To Amps** 10 milliamp = .010 amp We moved the decimal point 3 places to the left. 10. milliamp = .010 amp = .01 amp
Amps To Microamps 1 amp = 1000000 μamp We moved the decimal point 6 places to the right. 1 amp = 1.000000 amp 1.000000 amp = 1000000 μamp	**Microamps To Amps** 100000 μamps = .100000 amp We moved the decimal point 6 places to the left. 100000. μamp = .100000 amp = .1 amp

Fig. 3. Examples of conversion from one unit to another.

will find it is really quite simple and in a very short while you will find that you are converting from one unit to another mentally just as easily as you convert dollars and cents.

SUMMARY

In this section of the lesson you have reviewed many of the things you learned earlier about current flow. The important thing to remember about current flow is that it is a movement of electrons, and that in a series circuit the current flowing is the same at all parts of the circuit. Also remember that once the circuit is completed, current starts to flow in all parts of the circuit at the same instant.

You learned that to convert from rms values to peak values of ac current you multiply the rms value by 1.4 and to convert from peak values to rms values you divide by 1.4. The milliampere is one thousandth of an ampere. The microampere is one millionth of an ampere. To convert from the larger units to the smaller units you move the decimal point to the right and to convert from the smaller units to the larger units you move it to the left. In converting amperes to milliamperes and vice versa you move the decimal point three places, and in converting amperes and microamperes you move it six places. In converting microamperes and milliamperes you move the decimal point three places. These conversions are the same as converting dollars and cents and it is something that you will learn to do almost automatically.

Fig. 3 shows a number of examples of how to convert the different units. Be sure to study this figure carefully and then do the self-test questions. When you are doing these questions, if there is a conversion you do not know how to do, look at Fig. 3 and try to work it out yourself. If you can't, go to the back of the book and see how the conversion is made. Just look at the one you can't do and then go back to the self-test questions. There are several examples of each type of conversion. They are put in deliberately so that if you have trouble with the first one you can look at Fig. 3 to get help. Then look at the answer for additional help if necessary, and have another crack at doing it yourself. Changing back and forth becomes almost automatic - you simply move the decimal point the correct number of places.

SELF-TEST QUESTIONS

(a) If the current flowing past a point in the circuit is 1 ampere and it is increased so that four times the number of electrons pass the point in a second, what will the new current flow be?
(b) If the rms value of current is 3 amps, what will the peak value be?
(c) Change 7 amps rms to its peak value.
(d) If the peak value of current in a circuit is 7 amps, what is the rms value?
(e) In a circuit, the peak value of current is 21 amps; find the rms value.
(f) Convert $6.00 to cents.
(g) Convert 350 cents to dollars.
(h) Change 2 amps to milliamps.
(i) Convert 6 amps to milliamps.
(j) Convert 3.5 amps to milliamps.
(k) Convert .42 amps to milliamps.
(l) Convert .037 amps to milliamps.
(m) Convert .002 amps to milliamps.
(n) Convert 46 ma to amps.
(o) Convert 822 ma to amps.
(p) Convert 1327 ma to amps.
(q) Convert 2 amps to μa.
(r) Convert .0017 amps to μa.
(s) Convert 20 μa to amps.
(t) Convert 147 μa to amps.
(u) Convert .26 ma to μa.
(v) Convert .031 ma to μa.
(w) Convert 6100 μa to ma.
(x) Convert 927 μa to ma.
(y) Convert 327,000 μa to ma, and then to amps.

, 026 600

Voltage

You have already learned that voltage is the electrical pressure or force that can set electrons into motion. You know that two voltage sources are the battery and the generator. You also know that the unit in which voltage is measured is the volt.

Now let us review some of the important things you learned about voltage, and then expand that knowledge.

DC VOLTAGE

You know that dc is the abbreviation for direct current. A dc voltage is a voltage that will cause a direct current to flow. We refer to this voltage as dc voltage, not direct current voltage.

A battery is an excellent source of a dc voltage. The battery will supply a potential that will cause a constant current to flow in one direction. A dc voltage will cause a current flow from the negative terminal of the voltage source, through the circuit to the positive terminal of the source.

In the early days of radio, radios were built on a metal chassis and one side of the voltage source used to operate the radio was connected to this chassis. The chassis in turn was usually connected to a ground, such as a water pipe or a metal pipe driven into the ground. As a result, the connection to the chassis became known as the ground connection.

Today, you should not connect the chassis of a radio or television receiver to a ground connection because you might place a short across the power line if you do this. However, if the chassis is still connected to one side of the voltage source operating the radio or television set we still refer to it as a ground connection. Sometimes instead of using the chassis as a common connection for the various circuits in the receiver, we run a series of connections from one common point in the receiver to another and back to one side of the voltage source. We call this type of connection a "floating ground" to distinguish it from a ground that is actually connected to the chassis.

With one side of the voltage source connected to the ground, we make dc voltage measurements between the ground and various other points in the circuit. We then have a convenient method of indicating the voltage that should be found in various parts of the circuit.

An example of what this can lead to is shown in Fig. 4. Here we have shown a 15 volt battery connected to a small lamp. The positive terminal of the battery in Fig. 4A is numbered 1 and the negative terminal is numbered 2. Notice that connected to the negative terminal we have a lead going to a symbol that is marked ground. This is another new schematic symbol for you to remember.

We are going to start using an expression which may be new to you. It is "with respect to". It means simply "compared to" and is always used by electronics technicians. You will have no trouble if you know that "compared to" and "with respect to" mean the same thing.

In Fig. 4A the negative terminal of the battery is connected to ground. Using ground as a reference

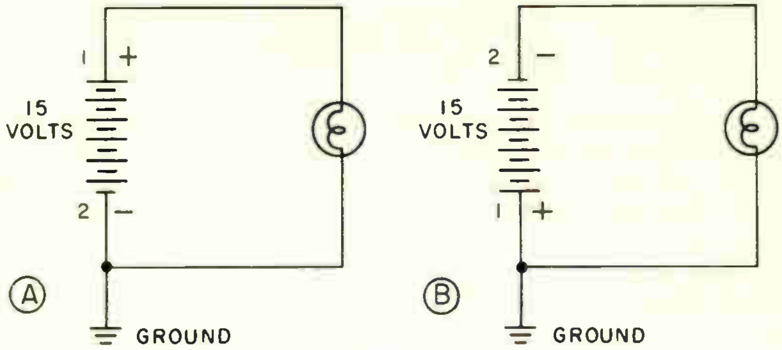


Fig. 4. Simple circuits with ground connections.

point we say that terminal 1 is +15 volts with respect to ground.

Now look at Fig. 4B. Here terminal 1 of the same battery, which is the positive terminal, is connected to ground. In this case using ground as a reference point we say that terminal 2 is -15 volts with respect to ground.

The important thing for you to understand in these two examples is that the voltage is the same in both cases. However, the polarity of the different points in the circuit will be positive or negative depending upon how the voltage source is connected to the circuit.

A 15 volt battery made up of 1-1/2 volt cells will have a total of ten cells. If the battery has a center connection, as shown in Fig. 5, and the center connection is connected to ground, we have a situation where we have both negative and positive polarities with respect to ground. Terminal 1 is + 7-1/2 volts with respect to ground and terminal 2 is - 7-1/2 volts with respect to ground. However, the total voltage is still 15 volts, as it was in both examples in Fig. 4, and the current flow through the bulbs would be the same in all three cases.

In your studies of electronic equipment you will run into equipment where the negative terminal of the voltage source is grounded as in Fig. 4A; you will run into equipment where the positive terminal is grounded as in Fig. 4B, and you will also encounter equipment where you have voltage both negative and positive with respect to ground as in Fig. 5.

AC VOLTAGE

You will remember that an ac voltage is one that causes an alternating current to flow. We refer to it as ac voltage instead of alternating current voltage. As in the case of dc voltage, we always use the abbreviation.

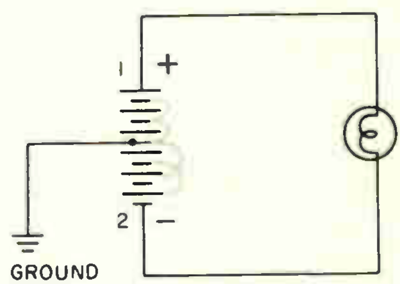


Fig. 5. A simple circuit where a voltage positive with respect to ground and a voltage negative with respect to ground are present.

Ac voltages are produced by generators or alternators. An alternator is a type of generator similar to the generator you studied earlier.

A schematic diagram of a simple ac circuit (similar to the dc circuit shown in Fig. 4) is shown in Fig. 6. Notice the symbol we have used to represent an ac generator. This is another important schematic symbol for you to remember.

In Fig. 6, terminal 2 of the generator is grounded. If we measure the voltage on terminal 1 with respect to ground or to terminal 2, we will get a voltage reading of 15 volts. The generator will produce an ac current flow that will have the same heating effect in the bulb as the 15 volt batteries did. However, you will remember that the voltage is continually varying. During one cycle, the voltage at terminal 1 starts at 0 as represented by point a in Fig. 7. The voltage begins to increase until it reaches its peak value at point b. If the rms or effective voltage is 15 volts then we know that the value of b will be 1.4 × 15 = 21 volts. Then the voltage between ground and terminal 1 begins to decrease until a half cycle after point a when we reach point c, where the voltage once again is 0. The voltage immediately begins to build up with the opposite polarity until we reach point d where once again the voltage between ground and terminal 1 is 1.4 × 15 = 21 volts. The voltage then begins to drop back until it reaches 0 again at point e.

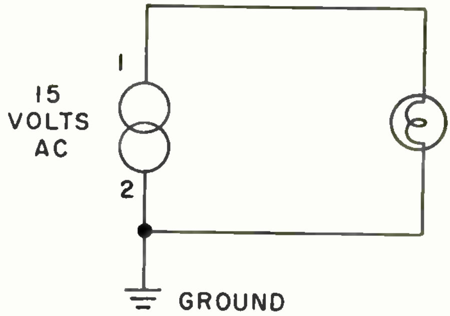


Fig. 6. A simple ac circuit with one side grounded.

The waveform between points a and e represents one complete cycle. The same cycle is completed again between points e and i, and then again between points i and m. In an ac circuit this cycle goes on indefinitely as long as the generator is operating.

During the first half cycle the voltage is positive at terminal 1 with respect to ground. During the next

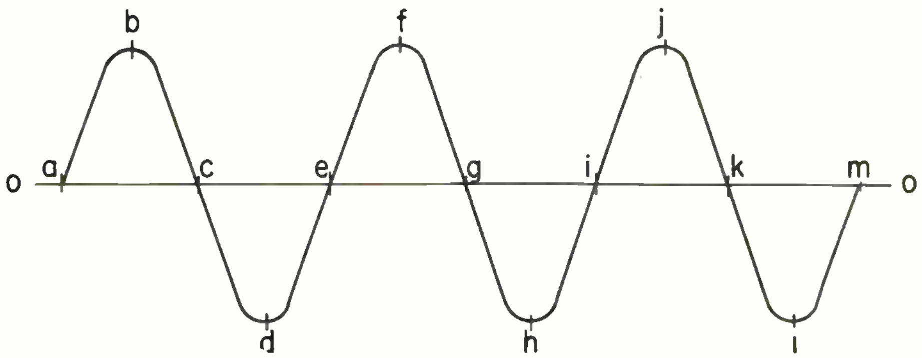


Fig. 7. Three cycles of an ac sinewave.

half cycle the voltage is negative with respect to ground. We draw the waveform above the line to represent a positive voltage and below the line to represent a negative voltage. We call the waveform between points a and b a quarter of a cycle; between b and c is also a quarter of a cycle as it is between c and d and d and e. The waveform between points a and c, points c and e, points e and g, etc., is referred to as a half cycle.

We said that we have a complete cycle between points a and e. By a complete cycle we mean the ac voltage starts at one point and goes through a complete cycle back to the equivalent point on the next cycle. We also have a complete cycle between points b and f because the waveform has gone through a full cycle between these two points. Between points c and g is also a full cycle as it is between points d and h. In speaking of a complete cycle we can start at any point on the cycle and continue on to the equivalent point on the next cycle. However, it is usually more convenient to start at a 0 point such as either point a or c when referring to a complete cycle. In fact, in most cases when speaking of an ac cycle we will start at point a and refer to the positive half of the cycle first and then the negative half. There is no reason why we have to do this; it is just something that is done by custom.

Notice the difference between the polarity at terminal 1 with respect to ground in Fig. 6 and the polarity of the ungrounded terminal in Figs. 4A and 4B. In Fig. 4A terminal 1 is the ungrounded terminal and it is always positive because the battery polarity does not change. In Fig. 4B, terminal 2 is the ungrounded terminal and it is always negative, again because the battery polarity does not change. In Fig. 6, terminal 1 is positive for one half cycle and negative for the next half cycle. Its polarity changes every half cycle because the voltage generated is an ac voltage. It is important for you to understand this difference between ac voltages and dc voltages.

VOLTAGES IN SERIES

Voltage sources can be connected in series. Whether they add together or subtract from each other depends upon the way in which they are connected.

In Fig. 8A we have shown two 4.5 volt batteries connected in series aiding. By this we mean that the two voltages add together. Notice that the batteries are connected together in the same way as the cells forming the battery are connected together. The positive terminal of the lower battery is connected to the negative terminal of the upper battery. With the arrangement shown in Fig. 8A the negative terminal of the lower battery is at ground potential. The voltage between terminal 1 and ground will be equal to the sum of the voltages of the two batteries, which in this case will be 9 volts.

In Fig. 8B, we have shown batteries with different voltages connected in series aiding. Notice that the positive terminal of the one battery is connected to the negative terminal of the other. The voltage between terminal 1 and ground will be the sum of the two battery voltages which is 7.5 volts. The position of the two batteries could be reversed; the voltage between terminal 1 and ground or terminal 2 would be the same in either case.

In Fig. 9, we have shown three examples of batteries connected in

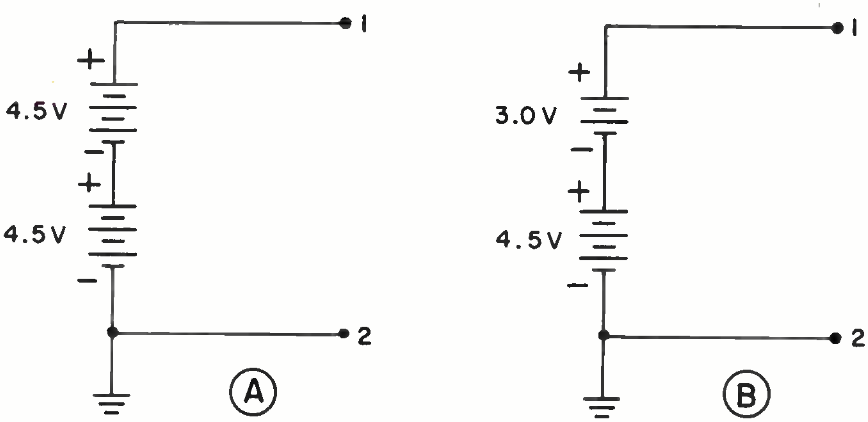


Fig. 8. Batteries connected in series so their voltages add.

series opposing. These batteries are connected so that their voltages oppose, and to find the total voltage we must subtract the battery voltages.

In Fig. 9A, each battery is a 4.5 volt battery. When you subtract 4.5 from 4.5 the result is 0 and therefore the potential between terminals 1 and 2 or between terminal 1 and ground is 0. The two batteries have equal voltages and therefore their voltages cancel.

In Fig. 9B, we have batteries of unequal voltages connected to oppose each other. The lower battery has a voltage of 3 volts, and the upper battery a voltage of 4.5 volts. Subtracting 3 from 4.5 gives us 1.5 volts. Since the upper battery has the higher potential, the voltage at terminal 1 will be positive with respect to ground or terminal 2. In other words, this voltage is able to overcome the voltage of the 3 volt battery and cause terminal 1 to be +1.5 volts with respect to ground.

In Fig. 9C, we have the opposite situation. Here the two batteries again subtract to give us a voltage

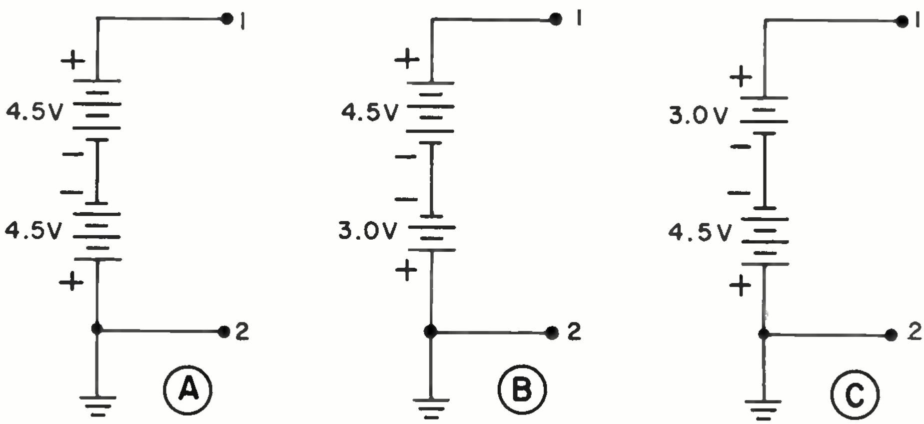


Fig. 9. Batteries connected in series so their voltages subtract.

of 1.5 volts. However, in this case, the polarity of the 4.5 volt and 3 volt batteries is reversed, so that now terminal 1 becomes -1.5 volts with respect to ground.

From the preceding examples you can see that batteries connected in series can either aid or oppose each other depending upon how they are connected. You will also see that when one side of the circuit is grounded, the other side can be either positive or negative depending upon the battery voltages and how they are connected. When unequal batteries are connected in series opposing, the polarity of the circuit will be the polarity of the battery with a higher voltage. When batteries are connected in series aiding, the polarity will be the same as the polarity of both batteries, since they must be connected in the same way in order to aid.

We can connect dc generators in series in exactly the same way as the batteries shown in Figs. 8 and 9 are connected. If they are connected in series aiding so that the negative terminal of one generator is connected to the positive terminal of the other, the total voltage available from the series combination will be the sum of the two voltages. If they are connected in series opposing so that the negative terminals of the generators or the positive terminals of the two generators are connected together, the voltage available will be the difference in voltage between the two generators, and the polarity of the circuit will be the polarity of the generator producing the higher voltage.

It is also possible to connect a battery in series with an ac generator such as shown in Fig. 10. Here, since the polarity of the voltage produced by the ac generator reverses every cycle we have a situation where during one half cycle the voltage produced by the generator will aid the battery voltage, and during the next half cycle the voltage will oppose the battery voltage.

In Fig. 11A, we have shown a graph of what the voltage will look like between terminals 1 and 2 of Fig. 10A when the peak voltage generated by the generator is exactly equal to the battery voltage. Terminal 2 is grounded and therefore is shown as zero voltage. The voltage at terminal 1 with respect to terminal 2 is shown

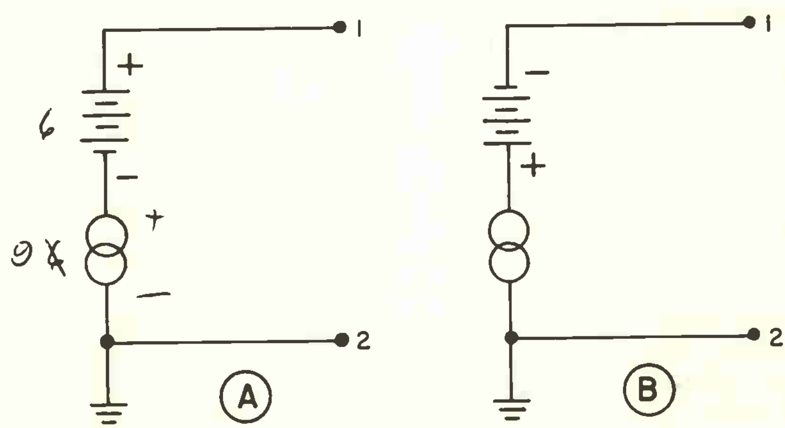


Fig. 10. Battery connected in series with ac generator.

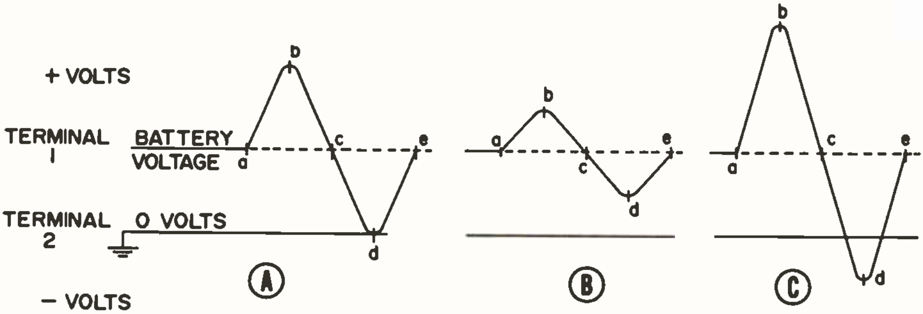


Fig. 11. Waveform showing how battery and generator voltages add in Fig. 10A. Voltage between terminal 1 and ground is shown.

by the graph. In electronics, as we have already pointed out, the expression "with respect to" means "compared to".

At the left of the graph in Fig. 11A, before the generator begins to turn, the voltage between terminals 1 and 2 will be the battery voltage. Terminal 1 is positive with respect to terminal 2 because the positive terminal of the battery is connected to terminal 1. When the generator begins to turn at point a in Fig. 11A, the generator voltage begins to build up during the first quarter cycle and adds to the battery voltage. When the generator voltage reaches its peak value at point b, the voltage between terminals 1 and 2 will be twice what it was at point a. This is because the peak generator voltage is adding to the battery voltage.

During the next quarter cycle, the generator voltage is falling from point b to point c. At the end of the first half cycle, the generator voltage will be zero as shown at point c, so the voltage between terminals 1 and 2 will be the battery voltage alone.

During the next quarter cycle, as shown between points c and d in Fig. 11A, the generator voltage begins to build up again, but this time the polarity of the generator voltage is reversed so it opposes the battery voltage. At point d, at the end of the third quarter cycle, the generator voltage will be exactly equal to the battery voltage but will have the opposite polarity. As a result, the two voltages will subtract, so at point d the voltage between terminals 1 and 2 will be zero. The actual voltage on terminal 1 at point d will be zero.

During the next quarter cycle, from point d to point e, the generator voltage decreases (becomes less negative) so that at the end of the cycle, at point e, the generator voltage is back to zero, and the voltage at terminal 1 is once again equal to the battery voltage.

In Fig. 11B, we have shown a situation in which the peak generator voltage is only half the battery voltage. Under these circumstances, the voltage builds up during the first quarter cycle. At point b, the total voltage between terminal 1 and terminal 2 will be 1-1/2 times the battery voltage. At point c on the curve, at the end of a half cycle, the generator voltage will be back to zero so the voltage between terminal 1 and ground (or terminal 2) will be equal to the battery voltage.

During the next half cycle, when the generator voltage begins to oppose the battery voltage, the voltage between terminals 1 and 2 drops until at the peak of the negative half cycle the voltage between terminals 1 and 2 is half the battery voltage. This is illustrated at point d in Fig. 11B.

As the ac voltage drops back to zero during the final quarter cycle, the voltage at terminal 1 increases back to the battery voltage as shown at point e.

In Fig. 11C, the peak generator voltage is 1-1/2 times the battery voltage. At the end of the first quarter cycle (at point b) the voltage between terminals 1 and 2 will be 2-1/2 times the battery voltage consisting of the generator voltage, which is 1-1/2 times the battery voltage, plus the battery voltage. At the end of the first half cycle (at point c) the generator voltage will be back to zero and the voltage between terminals 1 and 2 drops back to the battery voltage.

The peak generator voltage overcomes the battery voltage during the next half cycle because it is greater than the battery voltage and has the opposite polarity. As a result, at point d on the graph, terminal 1 is negative with respect to ground and terminal 2. Since the peak generator voltage is 1-1/2 times the battery voltage, it will cancel the battery voltage and then swing terminal 1 negative to a value equal to half the battery voltage.

You can see from Fig. 11 that when an ac voltage and a dc voltage are connected in series, they aid during one half cycle and oppose during the other half cycle. If the peak generator voltage is equal to the battery voltage, the total voltage will drop to zero volts once each cycle and will swing up to a value which is twice the battery voltage once each cycle. If the generator voltage is less than the battery voltage, the total voltage increases above the battery voltage and drops to less than the battery voltage once each cycle. On the other hand, if the generator voltage is greater than the battery voltage, the total voltage will reach a value which is more than double the battery voltage during one half cycle. When the generator voltage opposes the battery voltage during the other half cycle, the polarity of the output voltage will reverse when the generator voltage exceeds the battery voltage.

In Fig. 12, we have shown the voltage between terminal 1 and terminal 2, which is connected to ground, with the generator and battery connected as in Fig. 10B. The peak generator

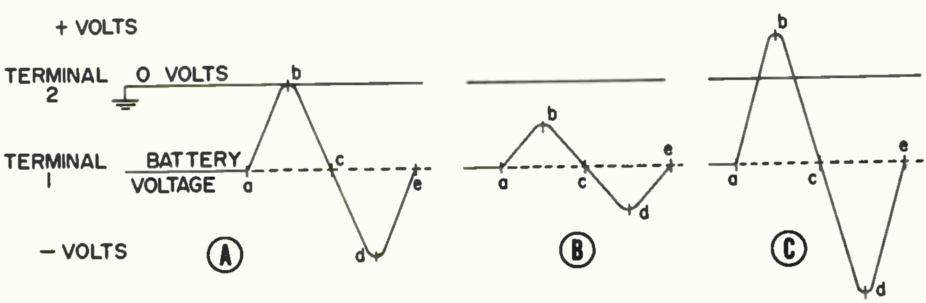


Fig. 12. Waveform showing how battery and generator voltages add in Fig. 10B. Voltage between terminal 1 and ground is shown.

voltage is exactly equal to the battery voltage in Fig. 12A, less than the battery voltage in Fig. 12B, and greater than the battery voltage in Fig. 12C.

When the peak generator voltage is exactly equal to the battery voltage, terminal 1 remains negative with respect to ground (terminal 2) at all times except when it drops to zero at the peak of the positive half cycle of the generator. When the peak generator voltage is less than the battery voltage as in Fig. 12B, terminal 1 remains negative with respect to ground (terminal 2) at all times. When the generator voltage is greater than the battery voltage as in Fig. 12C, terminal 1 becomes positive during a portion of the positive half cycle of the generator. For the remainder of the time, terminal 1 remains negative with respect to ground (terminal 2).

Having a battery and an ac generator connected in series is not much different from having two batteries connected in series except that in the case of the generator the polarity is changing each half cycle so that during one half cycle the two voltages aid and during the next half cycle the two voltages oppose. You can find the peak value that the two reach when they aid simply by adding the peak generator voltage to the battery voltage, and you can find the peak value that they reach when they oppose by subtracting the two. If the generator voltage is less than the battery voltage, then the voltage polarity in the circuit does not change. But if the generator voltage is greater than the battery voltage, the polarity of the voltage will change during the half cycle when the two voltages are opposing.

Why have we spent so much time explaining how voltages add or subtract? Do we ever make such connections in practical electronic circuits? This explanation was given to prepare you for actual circuits which will be described later. For example, in a tube or transistor we will have a fixed dc voltage so the part works properly. Then to this fixed dc voltage we will add a signal voltage, which is ac. The resulting combined voltage then becomes more negative or less negative (Fig. 12B) or more positive or less positive (Fig. 11B) and the tube or transistor can amplify the ac portion. This will be explained in greater detail later on. For the present, we just want you to know that there is a very practical reason for what you have just studied.

MILLIVOLTS

Just as the ampere is too large a current unit and we had to use milliamperes in some instances, so also is the volt sometimes too large a unit and we use the millivolt. A millivolt, abbreviated mv, is one thousandth of a volt. The prefix milli means the same when used with volts as it does with amperes - it means one thousandth. Therefore, to convert from volts to millivolts, you do the same thing as you did in converting from amps to milliamps. You multiply by 1000 - to do this you simply add zeros and move the decimal point three places to the right. To convert from millivolts to volts, you do the opposite; you divide by 1000 and to do this move the decimal point three places to the left. Remember, it is exactly the same as going back and forth between amps and milliamps. Thus, 2.5 volts = 2.5 × 1000 = 2500 millivolts; 49 millivolts = .049 volts.

MICROVOLTS

Sometimes even the millivolt is too large a unit and we use the microvolt, which is one millionth of a volt, just as the microampere is one millionth of an ampere.

To convert from volts to microvolts multiply by 1,000,000. You do this by moving the decimal point six places to the right. To convert from microvolts to volts you move the decimal point six places to the left, just as you did in converting microamps to amps.

You might wonder where such a small unit as the microvolt is used. It is sometimes used in measuring the strength of a radio or TV signal at a certain point. Also, in testing radio and television receivers, signals from a few microvolts are fed into the receiver and then measured at various points in the receiver to see how much the various stages are amplifying the signal. You might find that the signal at the input of one stage is 10 microvolts and at the output was 100 microvolts; this means that the stage amplified the signal voltage ten times. Technicians seldom have to convert back and forth between volts, millivolts and microvolts, as often as they do in the case of amps, milliamps and microamps. However, since the procedures for converting from one to the other are the same, if you know one, you know the other.

You will remember that to convert from milliamps to microamps you moved the decimal point three places to the right, and to convert from millivolts to microvolts you do the same thing. Similarly, to convert from microvolts to millivolts you move the decimal point three places to the left.

THE KILOVOLT

A unit that you will encounter in voltage measurements is the kilovolt. The kilovolt is one thousand volts. Thus 25 kilovolts, which is often abbreviated 25 kv, is equal to 25,000 volts. You will run into high voltages of this type in television receivers. Black and white television receivers often use voltages of 15 kv or more (15,000 volts) to operate the picture tube. In color television receivers some will have voltages as high as 25 kv. Just remember that a kilovolt is equal to 1000 volts, so to convert kilovolts to volts, you simply multiply by 1000.

SUMMARY

You should remember the important differences between ac and dc voltages. A dc voltage, which is produced by a battery or a dc generator, will have a polarity that does not change. Connected to a circuit, a dc voltage source produces a current which flows in one direction. An ac voltage changes potential (having one polarity during one half cycle and the opposite polarity during the next) and produces a current which flows in one direction during one half cycle and in the opposite direction during the next.

Voltages are often connected to common connections called grounds, and may be either positive or negative with respect to ground depending upon how they are connected to it.

Batteries and/or generators can be connected in series aiding so that their voltages add, or in series opposing so that they subtract. When connected in series between ground and another point, the polarity of the point may be positive or negative when the batteries oppose, depending on the polarity of the larger voltage source.

A dc voltage source such as a battery can be connected in series with an ac generator, and the two voltages will add during one half cycle and oppose during the next. The highest voltage produced is equal to the battery voltage plus the peak generator voltage. If one side of the circuit is connected to a common ground, the polarity in the circuit will not change unless the generator voltage is greater than the battery voltage. In this case, the polarity will change during the portion of the cycle when the voltages are opposing.

Some circuits you will encounter will have very small voltages; others will have very high voltages. Remember that the millivolt (abbreviated mv) is one thousandth of a volt, the microvolt (abbreviated μv) is one millionth of a volt, and the kilovolt (abbreviated kv) is one thousand volts.

SELF-TEST QUESTIONS

(z) If in the circuit shown in Fig. 4B, the battery is a 45 volt battery, what voltage is present at terminal 2 with respect to ground?

(aa) If in the circuit shown in Fig. 5, the battery is a 90 volt battery, and the ground terminal is a center tap, what is the voltage at terminal 1 with respect to ground? What is the voltage at terminal 2 with respect to ground?

(ab) In the circuit shown in Fig. 6, if the rms value of the voltage is 20 volts, what will the peak voltage be between terminal 1 and ground? Will it be positive or negative?

(ac) If a 15 volt battery and a 45 volt battery are connected in series as in Fig. 8B, what will the voltage be at terminal 1?

(ad) If two 22-1/2 volt batteries are connected in series as shown in Fig. 9A, what is the voltage at terminal 1 with respect to ground?

(ae) If two batteries are connected as shown in Fig. 9B, and one is a 45 volt battery and its positive terminal is connected to terminal 1, and the other is a 22-1/2 volt battery and its positive terminal is grounded, what will the voltage at terminal 1 be?

+ 45v − − 22.5v +

(af) If in Fig. 9C the positions of the two batteries are reversed so that the negative terminal of the 4.5 volt battery is connected to terminal 1, and the negative terminal of the 3 volt battery connects to ground, what will the voltage be at terminal 1?

(ag) If a generator that has a peak voltage of 15 volts is connected in series with a battery that has a voltage of 15 volts, what is the maximum voltage produced during the half cycle when the two aid, and what will be the minimum voltage produced when the two oppose?

(ah) If in the circuit shown in Fig. 10A the battery voltage is 15 volts and the peak generator voltage is 20 volts, what will the voltage be between terminal 1 and ground when they are aiding and when they are opposing?

(ai) If a generator and battery are connected as in Fig. 10B, and the battery voltage is 30 volts and the generator peak voltage is 45 volts, what will the voltage be between terminal 1 and ground when the two are aiding and when the two are opposing?

(aj) If a battery and generator are connected as shown in Fig. 10A and the battery voltage is 20 volts and the peak generator voltage is 10 volts what will the voltage be between terminal 1 and ground when they are aiding? When they are opposing?

(ak) If a battery and generator are connected in series as shown in Fig. 10B, and the battery voltage is 45 volts and the peak generator voltage is 30 volts, what will the voltage be between terminal 1 and ground when the generator reaches its peak value aiding the battery and when it reaches its peak value opposing the battery?

Resistance

One of the most important values you will work with in your electronic career is resistance. All wires and parts in electronic equipment have a certain amount of resistance. In some cases, such as in a short piece of copper wire, the resistance may be so low that it has no effect on the performance of the circuit. However, in every circuit there will be some part that has enough resistance to affect the operation of the circuit.

You know that when a voltage is applied to an electrical circuit a current will flow in the circuit. In a dc circuit, the thing that limits the amount of current that will flow for a given voltage is the resistance of the circuit. In an ac circuit, resistance also limits the current flow, but there may also be some other parts that will affect the current flow.

THE OHM

The unit of resistance is the ohm. It is named after the scientist George Simon Ohm who did a great deal of work in the early days when scientists first began studying electricity. If a voltage of 1 volt is applied to a circuit and a current of 1 amp flows in the circuit, the resistance of the circuit is 1 ohm. If a voltage of 2 volts is applied to a circuit, and a current of 1 amp flows, the resistance in the circuit is 2 ohms. Here we have twice the voltage applied to the circuit and, therefore, twice the force to force a current flow in the circuit. However, since the current flow is 1 amp in both cases, the circuit in the second case must offer twice the opposition to current flow. This is why the resistance of the circuit is 2 ohms.

In electronic equipment, the wires used to connect parts together have a very low resistance, usually only a fraction of an ohm. However, several of the parts that you have studied do have a much higher resistance. For example, a transformer has two or more windings on a common core. In the case of a transformer used to operate on a 60-cycle power line, there will be many turns on the primary winding of the transformer. The resistance will depend upon the size of the wire used on the primary winding of the transformer, but a resistance of about 100 ohms is typical of what you might find if you measured the resistance of the primary winding of the transformer used in a large radio or a small television receiver.

DC Resistance.

Dc resistance is the opposition offered to the flow of direct current in a circuit. If a dc voltage is applied to an electrical circuit, the dc resistance of the circuit will limit the current that will flow in the circuit. When we speak of dc resistance, we simply refer to it as resistance rather than by its entire name "dc resistance".

AC Resistance.

The ac resistance of a part may not be the same as the dc resistance. For example, at high radio frequencies, which are simply very high ac frequencies, the current flowing in a circuit has a tendency to flow on the outside of the conductor. This often causes the ac resistance to be somewhat higher than the dc resistance. In some coils used in very high frequency equipment you will find that the coils are silver plated. The purpose of the silver plating is to keep the resistance on the outside of the conductor as low as possible.

At power line frequencies and audio frequencies as well as low radio frequencies, the ac resistance of most parts is almost the same as the dc resistance and so we generally consider them as being the same. You will see later that it is comparatively easy to measure the dc resistance of a part, but it is much more difficult to measure the ac resistance.

Resistors.

While copper wire is used to connect electronic parts together to keep the resistance in the circuit low, there are some instances where we want resistance in the circuit. Parts made to put resistance in the circuit are called resistors. There are many different values and sizes of resistors used in electronic equipment, and several different types, but the most commonly used type is the "carbon resistor". This type of resistor is made of a mixture of powdered carbon and a cement-like material that is used to hold the carbon together. By varying the composition of the mixture, different values of resistance from a few ohms up to several million ohms can be obtained.

Carbon resistors come in several different sizes and in many different resistance values. The size of the resistor tells you how much power the resistor can handle. Three different sizes of carbon resistors are shown in Fig. 13. Each resistor has a resistance of 1000 ohms. The resistor in the middle can handle twice the power that the resistor on the top can handle. The resistor on the bottom can handle twice the power the resistor in the middle can, or four times the power the resistor

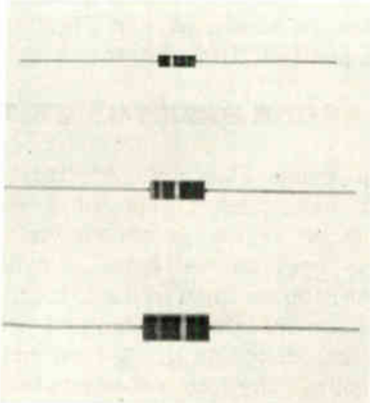


Fig. 13. Three 1000-ohm carbon resistors. The resistor at the top is a ½ watt size, the middle resistor, a 1 watt size and the bottom one, a 2 watt resistor.

on the top can handle. The resistor on the top is called a half-watt resistor, the one in the middle a one-watt resistor and the one on the bottom a two-watt resistor. The watt is a unit of electrical power, which you will learn about later.

Another type of resistor that you will encounter is the "wire-wound resistor". It is made of wire wound on a form, which is usually some type of ceramic form. The wire used to wind the resistor is called resistance wire; it gets its name because it has a much higher resistance than copper wire. Wire-wound resistors are used in places where they must handle a higher current than could be handled by a carbon resistor.

Another type of resistor that you will encounter is the deposited film resistor. This type of resistor has a metal oxide (a combination of a metal and oxygen) film deposited on a ceramic form. The advantage of this type of resistor is that it can be made to handle higher currents than a carbon resistor and at the same time can be made in larger resistance values than the wire-wound resistor. A wire-wound resistor is shown at A in Fig. 14 and a deposited film resistor at B.

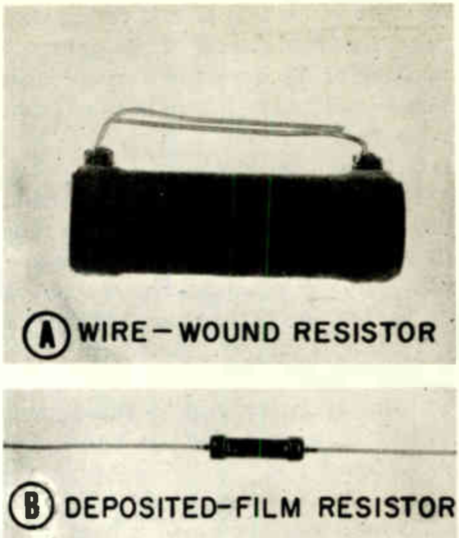


Fig. 14. A wire-wound resistor is shown at A; a deposited film resistor at B.

LARGER RESISTOR UNITS

In many electronic circuits you will have resistances of several thousand ohms; in others you will have resistances over a million ohms. Rather than indicate the value of these resistors in ohms it is more convenient to use the K-ohm and the megohm. The letter K stands for one thousand so the K-ohm is one thousand ohms. Meg stands for one million so one megohm is one million ohms. Thus, rather than mark the value of a resistance as 2,200 ohms we would indicate the value as 2.2K. A resistor having a resistance of 100,000 ohms would be labelled 100K. A resistor with a resistance of 470,000 ohms would be marked 470K.

We use the unit megohms for resistors larger than one million ohms. A resistor whose value is 2,200,000 ohms would be labelled 2.2 megs or 2.2M; both abbreviations are used. Sometimes a resistor that is somewhat less than a megohm in resistance is also expressed in megohms. For example, a 470,000 ohm resistor could be labelled 470K, and it can also be labelled .47 megs or .47M. Any one of the three labellings could be used since they all mean the same thing.

Converting back and forth between ohms, K-ohms and megohms is essentially the same as converting between amps, milliamps and microamps. However, in this case remember that the ohm is the small unit, the K-ohm is one thousand ohms and the megohm is one million ohms. To convert from the small unit to the larger unit you simply move the decimal point to the left, either three places or six places, depending on whether you are converting to K-ohms or to megohms. To convert from the large unit to the small unit you move the decimal point in the opposite direction. As a technician you will have to convert values back and forth. The values of carbon resistors are identified by means of a color code. The color code will give the resistance in ohms, but the value may be given on a circuit diagram in K-ohms or megohms to save space. Thus you have to know what the different units mean so you will be able to identify them on circuit diagrams.

SUMMARY

Much of the material covered in this section of the lesson will be a review for you. However, resistance is a very important subject, and you should make sure you understand everything in this section before going on to the next. In the next section of the lesson you are going to study Ohm's Law.

The important points to remember in this section are that in a dc circuit the current flow in the circuit will be limited by the resistance in the circuit. In an ac circuit the resistance will also limit the current, but there may be some other factors that also aid in limiting the current.

The unit of resistance is the ohm. If a current of 1 ampere flows in a circuit when a voltage of 1 volt is applied to the circuit, the resistance in the circuit is 1 ohm.

Three important types of resistors that you will encounter in electronic equipment are the carbon resistor, the wire-wound resistor, and the metal oxide film resistor. These resistors are made in many different resistance values and in different sizes to handle different values of current.

In many electronic circuits the resistance is so high that we use the K-ohm, which is equal to one thousand ohms and the megohm which is equal to one million ohms. You should be able to convert from one unit to another so you will be familiar with all three units.

SELF-TEST QUESTIONS

(al) If the current flowing in a circuit is 1 amp, and we double the resistance in the circuit, will the current increase, decrease, or remain the same?
(am) Name the three types of resissistors that are used in electronic equipment.
(an) Convert 4,700 ohms to K-ohms.
(ao) Convert 5,600,000 to megohms.
(ap) Convert .330 megs to K-ohms.
(aq) Convert 2.2 megs to ohms.
(ar) Convert 8.2 K-ohms to ohms.
(as) Convert 680 K-ohms to megohms.
(at) Draw the symbol used to represent a resistance. You should remember this from an earlier lesson.

Ohm's Law

Ohm's Law is one of the most important laws or rules in electronics. It tells you how the voltage, current and resistance are related in an electrical circuit. Ohm's Law states that the current flowing in the circuit is equal to the voltage divided by the resistance. Rather than use words every time to express this law, we use symbols. We use the letter I for current, E for voltage and R for resistance. Using these symbols we can express Ohm's Law as:

$$I = E + R$$

This is more often written in the form:

$$I = \frac{E}{R}$$

By using this expression, or by rearranging it mathematically, if we know any two of the three values, resistance, current or voltage, we can determine the other. As a radio television serviceman you will be concerned with replacing parts and will not have many occasions to work out the value of a part using Ohm's Law. However, an understanding of it will help you understand what is going on in the circuits you will study in this and in following lessons. A technician who wants to get his FCC license to work at a broadcasting or television station, or a technician who wants to work in industry or as an engineering aid will have to be able to work out problems involving Ohm's Law. We are going to use Ohm's Law to learn more about how voltage and resistance affect the current in a circuit. We will do some simple problems involving Ohm's Law. Be sure to follow these through carefully so you will understand what is happening in the circuits we discuss. We will work out each step in detail even though some of the steps may seem comparatively simple.

HOW VOLTAGE AFFECTS CURRENT

In Fig. 15 we have shown a simple circuit consisting of a voltage source

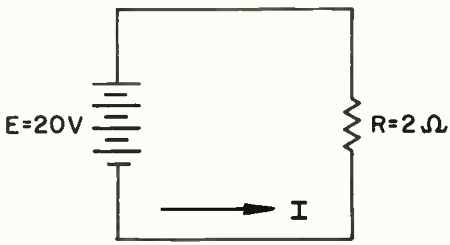


Fig. 15. A simple circuit consisting of a voltage source and a resistor.

and a resistor. The battery voltage, E = 20 volts; the resistance of the resistor, R = 2 ohms. Notice the symbol we have used to indicate ohms. This is the Greek letter omega and it is often used on diagrams as an abbreviation for ohms.

We can determine the current that will flow in this circuit by using Ohm's Law. We simply take the formula and then substitute the values of E and R which we have and this will give us the value of I.

$$I = \frac{E}{R}$$

$$I = \frac{20}{2} = 10 \text{ amps}$$

Thus in the circuit shown in Fig. 15 the current flowing will be equal to 10 amps. If we increase the voltage to 40 volts, the new current flowing will be:

$$I = \frac{40}{2} = 20 \text{ amps}$$

and if we reduce the voltage to 10 volts, the current flowing in the circuit will be:

$$I = \frac{10}{2} = 5 \text{ amps.}$$

The important thing to see from the preceding is how the current flowing in the circuit with a given resistance in the circuit is tied directly to the voltage. Increasing the voltage increased the current, and reducing the voltage reduced the current. In the example where we doubled the voltage, the current doubled, and where we cut the voltage in half, the current was cut in half. This relationship will always hold true. If we increase the voltage to three times its original value, then the current will increase to three times the original value, and if we reduce the voltage to one third of its original value, then the current will be reduced to one third of its original value. We can say that in a given circuit the current will vary directly with the voltage. Any change in the voltage will result in a corresponding change in the current.

Notice that in the example given, the voltage was given in volts, the

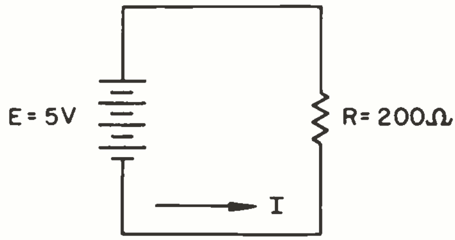


Fig. 16. A simple circuit where E = 5V and R = 200Ω.

resistance was given in ohms and the current we calculated was in amps. In using Ohm's Law we must always use these basic units. Often this leads to a somewhat more difficult problem than shown in Fig. 15. Look at the example in Fig. 16.

In Fig. 16, the voltage is 5 volts and the resistance in the circuit is 200 ohms. Using Ohm's Law to determine the current we have:

$$I = \frac{5}{200}$$

Here the division is not quite as simple as it was in the preceding example because we have to resort to a decimal division. The division is not particularly difficult; it is shown in Fig. 17. We see that we get a current of .025 amps. We could leave our answer like this, or we can convert it to milliamperes by multiplying by 1000. To do this we simply move the decimal point three places to the right and get the answer, I = 25 ma.

```
        .025
 200 ) 5.000
       4 00
       1 000
       1 000
```

Fig. 17. Solution for I in Fig. 16.

Since our answer is in milliamperes, and we multiplied the current in amperes by 1000 to get our answer into milliamperes, if we want to avoid the decimal division we can multiply by 1000 before performing the division and get the answer directly in milliamperes. When we do this the problem becomes:

$$I = \frac{5}{200} \times 1000$$

$$I = \frac{5000}{200}$$

Now we can cancel two zeros above the division line and two below the line so we have:

$$I = \frac{50\cancel{0}\cancel{0}}{2\cancel{0}\cancel{0}} = \frac{50}{2}$$

$$I = 25 \text{ ma}$$

You can use whichever method you want in determining the current in a circuit of this type. If you multiply by 1000 before performing the Ohm's Law division, be sure to remember that your answer will be in milliamperes. Incidentally, sometimes we use mils as well as ma for an abbreviation of milliamperes. In the preceding problem we can say our answer is 25 mils.

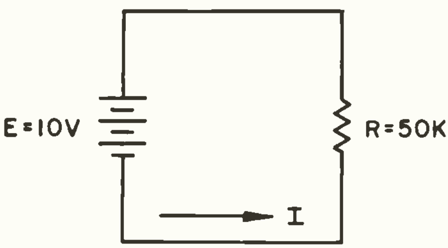


Fig. 18. A simple circuit where E = 10V and R = 50 K-ohms.

Another example is shown in Fig. 18. Notice that here we have an even larger resistor and also notice that the resistance is given in K-ohms. Let us see how we go about tackling a problem of this type.

You will remember we said that in using Ohm's Law the units must be in volts, amps and ohms. Therefore, the first thing we must do is to convert 50K-ohms to ohms. We do this simply by multiplying by 1000 and get a resistance of 50,000 ohms. Now using 50,000 ohms and 10 volts we can find the current using Ohm's Law.

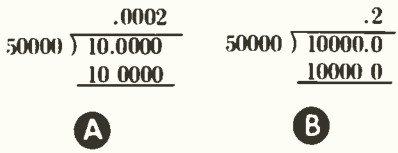


Fig. 19. Solution for I in Fig. 18.

$$I = \frac{E}{R}$$

$$I = \frac{10}{50,000}$$

Here again we have a decimal division. We can go ahead and perform this division as shown in Fig. 19A. We get as an answer .0002 amps. We can convert this to milliamperes by multiplying by 1000. To do this, we move the decimal point three places to the right and get .2 milliamps as the current. However, we still have a decimal so instead of converting to milliamps it would be more logical to convert to microamps. We do this by multiplying by 1,000,000 and this involves moving the decimal point six places to the right. We get as the current 200 μa.

If we want to avoid the decimal

division we can again convert before we perform the division. If we try converting to milliamps first by multiplying by 1000 we would have 50,000 divided into 10,000. This can be worked out as shown in Fig. 19B and we get our answer, .2 ma. However, since 50,000 is larger than 10,000 it is obvious that we have another decimal division and so to avoid this why not simply multiply the voltage, 10 volts by 1,000,000, and then get our answer directly in microamps. When we do this the formula becomes

$$I = \frac{10}{50,000} \times 1,000,000$$

$$= \frac{10,000,000}{50,000}$$

Notice that in 50,000 there are four zeros so we can mark these four zeros off and mark four zeros off from the ten million so that we will have

$$I = \frac{10,00\cancel{0},\cancel{0}\cancel{0}\cancel{0}}{5\cancel{0},\cancel{0}\cancel{0}\cancel{0}}$$

$$= \frac{1,000}{5}$$

$$= 200 \text{ microamps}$$

In electronic circuits you will frequently run into comparatively small voltages and very large resistances. This means that if you want to find the current flowing in the circuit and you divide directly, you will run into a decimal division. Therefore the easy way is to multiply by 1,000 or 1,000,000 first to convert the answer directly to milliamps or microamps. If you are doubtful about whether you should multiply by 1000 or 1,000,000, multiply by 1,000,000 and then perform the division. If your answer, which will be in microamperes, is over 1,000 microamperes you can convert this to milliamps if you want to by simply moving the decimal point three places to the left. However, it really doesn't matter whether you say that the current is 4,700 microamps or 4.7 milliamps, it means the same thing. For that matter, if you don't mind doing decimal divisions you don't have to convert at all. You can simply give the answer in amperes as .0047 amps. All three mean the same thing. If you like doing mathematics then the chances are that the decimal divisions won't bother you, but on the other hand if you are like most people and steer away from math, then multiplying either by 1,000 or a 1,000,000 first to convert the answer directly to milliamperes or microamperes is probably the easiest way to tackle problems of this type.

In examples shown in Figs. 15, 16 and 18, the voltage source was a battery and therefore the problems all involved dc voltage and current. If the voltage source had been an ac generator instead, we would have found the current in exactly the same way. In a circuit containing only resistance, Ohm's Law is used in exactly the same way in an ac circuit as it is in a dc circuit. Where the value of E is the rms value of the voltage, then the current will be found in its rms value. If the value of E given is the peak value of the voltage, and we use this value in Ohm's Law, then we will get the peak value of the ac current. If we want the rms value we can get it either by converting the peak value of the current to its rms value after we perform the calculation or we can convert the peak voltage to an rms value first.

HOW RESISTANCE AFFECTS CURRENT

In the simple circuit shown in Fig. 15 where the voltage is 20 volts and the resistance 2 ohms, we found by using Ohm's Law that the current flowing in the circuit is 10 amps. The same circuit is repeated in Fig. 20 except that we have replaced the 2 ohm resistor with a 4 ohm resistor.

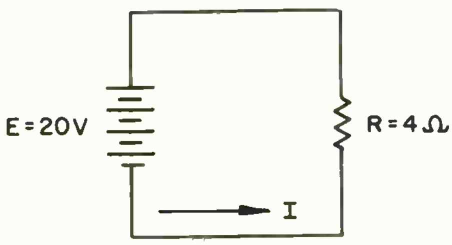


Fig. 20. The 2Ω resistor in the circuit of Fig. 15 is replaced by a 4Ω resistor.

How will this affect the value of I? Using Ohm's Law we can find the current.

$$I = \frac{E}{R}$$

$$I = \frac{20}{4} = 5 \text{ amps}$$

Here the current is 5 amps, half of what it was before. In other words, doubling the resistance cut the current in half. If instead of doubling the resistance, we had cut it in half so that R is equal to 1 ohm, then using Ohm's Law again we find that the current is equal to 20 amps. In other words, cutting the resistance in half has doubled the current. We will find that this relationship between current and resistance holds true regardless of how we change the resistance. If we increase the resistance to 3 times its original value, the current will be reduced to one third and if we cut the resistance to one third its original value the current will increase to three times its original value. We say that the current varies inversely to the resistance.

As a matter of fact, this relationship between current and resistance is obvious if we examine Ohm's Law. If we look at the expression for current

$$I = \frac{E}{R}$$

if E remains constant and we increase R, it is obvious that I must be smaller. Similarly, if we reduce R, and keep E constant, then I must get larger.

FINDING E

In some circuits we may know what the current flowing in the circuit and the resistance in the circuit are and have to find the voltage in the circuit. An example of this type of problem is shown in Fig. 21. Here the current is 2 amps and the resistance is 15 ohms. We want to find the value of E.

Ohm's Law can be rearranged mathematically into the form

$$E = I \times R$$

We usually drop the multiplication sign and simple write the formula as:

$$E = IR$$

The term IR means I × R; the multiplication sign is understood, even though it is not shown. Using this form of Ohm's Law we get:

$$\begin{aligned} E &= 2 \times 15 \\ &= 30 \text{ volts} \end{aligned}$$

Thus, the value of the voltage applied to the circuit must be 30 volts. If we want to, we can check this out using the other form of Ohm's Law:

$$I = \frac{E}{R}$$

substituting 30 for E and 15 for R we get a current of 2 amps, and since this agrees with the value given, the value of voltage which we determined must be correct.

Notice that the current in Fig. 21 is given in amps and the resistance in ohms. Sometimes the current may be given in milliamperes or microamperes and the resistance might be in K-ohms or megohms. In either case we must convert back to the basic units. Milliamperes and microamperes must be converted to amperes and K-ohms and megohms must be converted to ohms. An example of this type of problem is shown in Fig. 22. Here the current is given in milliamperes and the resistance in ohms.

Doing the problem shown in Fig.

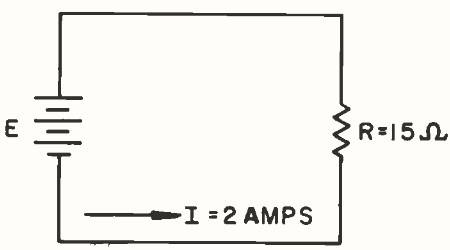


Fig. 21. A simple circuit where I = 2 amps and R = 15Ω.

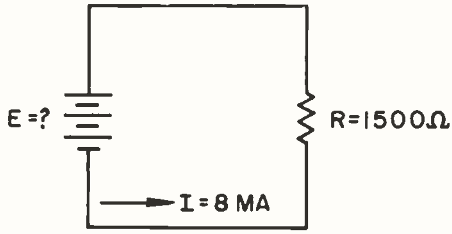


Fig. 22. Find E where I = 8 ma and R = 1500Ω.

22, we can first convert the current I which is 8 ma to amps by dividing by 1000. We do this by moving the decimal point three places to the left and get as the current, .008 amps. Now substituting these values in Ohm's Law we can find the voltage

$$\begin{aligned} E &= IR \\ &= .008 \times 1500 \\ &= 12 \text{ volts} \end{aligned}$$

This involves a decimal multiplication, which is not too difficult to perform. However, if you want to avoid the decimal multiplication the easy way to do this is to simply write 8 milliamps as 8/1,000 amps. Now substituting the value of 8/1,000 in the formula we get

$$E = \frac{8}{1000} \times 1500$$

$$E = 8 \times \frac{1,500}{1,000}$$

$$E = \frac{12,000}{1,000}$$

Now cancelling three zeros above and below the division line we get

$$E = \frac{12,\cancel{0}\cancel{0}\cancel{0}}{1,\cancel{0}\cancel{0}\cancel{0}}$$

$$E = 12 \text{ volts}$$

If in Fig. 22 the resistance had been expressed in K-ohms it would be written 1.5 K-ohms. To convert this to ohms we multiply by 1,000. Instead of actually performing the multiplication we could then write the problem as:

$$E = \frac{8}{1{,}000} \times 1.5 \times 1{,}000$$

Now we can cancel the 1,000 above and below the division line and simply multiply 8 × 1.5 and our answer is 12 volts.

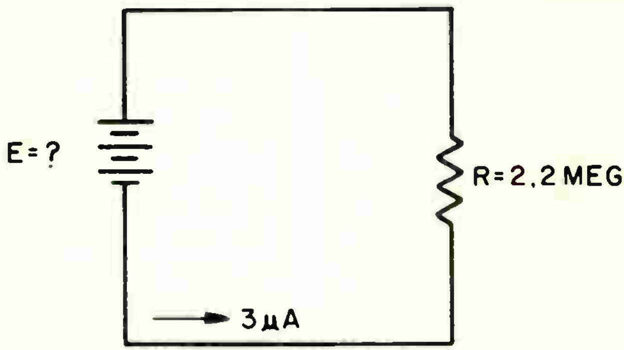


Fig. 23. Find E where I = 3 μa and R = 2.2 megohms.

In the example shown in Fig. 23 the current is given in microamps and the resistance in megohms and we have to find the voltage. We can convert microamperes to amperes by dividing the current by 1,000,000 and megohms to ohms by multiplying the resistance by 1,000,000. Thus substituting the formula:

$$E = I \times R$$

$$E = \frac{3}{1{,}000{,}000} \times 2.2 \times 1{,}000{,}000$$

$$= \frac{3 \times 2.2 \times 1{,}000{,}000}{1{,}000{,}000}$$

Here we simply cancel the 1,000,000 above and below the division line and multiply 3 × 2.2 and get our answer 6.6 volts. It is far simpler to do the problem this way than it is to convert to amps and ohms. These same techniques can be used regardless of what units the current and resistance are in. Remember that if the current is given in milliamperes simply write it over 1,000 and this will convert it to amps. If it is given in microamperes write it over 1,000,000 and this will convert it to amps. If the resistance is given in K-ohms multiply it by 1,000 to convert it to ohms and if it is given in megohms multiply it by 1,000,000 to convert it to ohms. Then before doing anything else look for numbers above and below the division line that can be cancelled and then perform the remaining multiplication.

FINDING R

We can rearrange Ohm's Law to find the resistance in a circuit if we know the voltage and the current. To do this we use Ohm's Law in the form:

$$R = \frac{E}{I}$$

Let us go back to the circuit shown in Fig. 15. Here the voltage is 20 volts and we found that the current is 10 amps. Now let us use Ohm's

Law to see what value of resistance we get using these values of voltage and current.

$$R = \frac{E}{I}$$

$$R = \frac{20}{10}$$

$$= 2 \text{ ohms}$$

As you see, the value obtained for the resistance is the value given originally so that we know that our calculation is correct.

When the voltage is given in volts and the current given in amps, determining the resistance in the circuit is quite simple. However, often the current will be in milliamps or microamps. In this case we must convert the current to amps in order to use the formula.

To convert milliamps to amps we divide by 1,000. We can do the same thing simply by multiplying the voltage by 1,000 and using the current in milliamperes in the formula. Thus the formula for resistance becomes

$$R = \frac{E \times 1{,}000}{I \text{ (in ma)}}$$

If the current is given in microamps we can multiply the voltage by 1,000,000 and then use the current in microamperes in the formula. Then the formula becomes

$$R = \frac{E \times 1{,}000{,}000}{I \text{ (in } \mu a)}$$

in microamperes.

Multiplying the voltage by 1,000 when the current is in milliamperes and by 1,000,000 when it is in microamperes eliminates the necessity of performing a decimal division. It is usually easier to divide by whole numbers and then multiply the result by either 1,000 or 1,000,000 than it is to perform the conversion to amperes first and then perform a decimal division. An example of a problem where the current is in milliamperes is as follows: find the resistance in a circuit where the voltage is 3 volts and the current is 3 milliamperes. Using the formula:

$$R = \frac{E \times 1{,}000}{I \text{ (ma)}}$$

$$R = \frac{3 \times 1000}{3}$$

$$R = 1000 \text{ ohms}$$

An example where the current is in microamperes is: find the resistance in a circuit where the voltage is 8 volts and the current 100 microamperes.

Using the formula:

$$R = \frac{E \times 1{,}000{,}000}{I \text{ } (\mu a)}$$

$$R = \frac{8 \times 1{,}000{,}000}{100}$$

the easiest way to do this problem is to divide the top and bottom by 100 simply by cancelling the 100 on the bottom and removing two zeros from the 1,000,000 and then we have

$$R = 8 \times 10{,}000$$
$$= 80{,}000 \text{ ohms}$$

SUMMARY

In this section of this lesson you have seen how the current in a circuit is affected by the voltage and the resistance in the circuit. We found that increasing the voltage increased the current and decreasing the voltage decreased the current. We said that the current varies directly as the voltage. In the case of the resistance in the circuit we found it had the opposite effect on the current. Increasing the resistance in the circuit decreases the current and decreasing the resistance increases the current. We say that the current varies inversely with the resistance.

We saw the three important forms of Ohm's Law and how you can use it in solving problems involving voltage, current and resistance in a circuit. If any two of these three quantities are known, you can use Ohm's Law to find the other.

You should remember the three forms of Ohm's Law. You will use them over and over again, so it would be worthwhile to take time now to memorize them. The three forms are:

$$I = \frac{E}{R}$$

$$E = IR$$

$$R = \frac{E}{I}$$

Remember that to use Ohm's Law the voltage must be in volts, the current in amps and the resistance in ohms. We showed you simple ways of getting around the problem of performing decimal operations in each type of problem. As you do the self-test questions do not hesitate to go back to the section of the lesson that dealt with the particular kind of problem you are working on and review how we worked the problem. It is not so important for you to memorize how to do these problems as it is to be able to do them with the aid of the examples given in the textbook. You will not be able to remember all you will learn about electronics either during your course or after you have completed your course, but the important thing is to remember the basic fundamentals and then know where to find the other facts that you may need.

If you find that you are puzzled by one of the self-test questions and you can't work it out even after reviewing the lesson, find the answer to the question at the back of the book and see how we did the problem. Then close the book and try to do the problem yourself and other problems of the same type without referring to the answers again. Don't be discouraged if you don't get the right answer every time or if you have trouble remembering the various forms of Ohm's Law at first. Make a determined effort to memorize them and after you have used them a number of times you will find that they will remain in your mind.

SELF-TEST QUESTIONS

(au) Give the form of Ohm's Law that is used when you know the voltage and the current and want to find the resistance.

(av) Write the Ohm's Law formula that is used when you know the current and resistance in a circuit and want to find the voltage.

(aw) Write the Ohm's Law formula used when you know the voltage and resistance in a circuit and need to find the current.

(ax) In a circuit such as Fig. 15, find the current if the voltage is 15 volts and the resistance is 3 ohms.

(ay) In a circuit like Fig. 15, find

the current if the voltage is 12 volts and the resistance is 6,000 ohms.

(az) In a circuit like Fig. 15 find the current if the voltage is 28 volts and the resistance 7,000 ohms.

(ba) In a circuit like Fig. 15 find the current if the voltage is 15 volts and the resistance 300 K-ohms.

(bb) Find the voltage in a circuit like the one in Fig. 21 where the current is 3 amps and the resistance is 6 ohms.

(bc) Find the voltage in a circuit like the one shown in Fig. 21 where the current is 20 ma and the resistance 1,000 ohms.

(bd) Find the voltage in a circuit like the one shown in Fig. 21 when the current is 18 ma and the resistance is 3K-ohms.

(be) Find the voltage in a circuit like Fig. 21 when the current is 47 μa and the resistance 200,000 ohms.

(bf) Find the voltage in a circuit like Fig. 21 when the current is 58 μa and the resistance 330K-ohms.

(bg) Find the voltage in a circuit like Fig. 21 when the current is 6 μa and the resistance 2 megohms.

(bh) In a simple series circuit if the applied voltage is 12 volts, and the current is 4 amps, what is the resistance in the circuit?

(bi) Find the resistance in the circuit if the applied voltage is 24 volts and the current 8 ma.

(bj) Find the resistance in the circuit if the applied voltage is 45 volts and the current flowing is 15 ma.

(bk) If the voltage applied to a circuit is 34 volts, and the current flowing is 170 μa, what is the resistance in the circuit?

(bl) If the voltage applied to a circuit is 144 volts, and the current flowing in the circuit is 120 μa, what is the resistance in the circuit?

(bm) Fill in the missing words in the following statement: if the voltage applied to a circuit is increased, the current will ____, and if the voltage applied to a circuit is decreased, the current will __________.

(bn) Fill in the missing words in the following: if the resistance in a circuit is increased, the current will _____, and if the resistance in a circuit is decreased, the current will ____.

ANSWERS TO SELF-TEST QUESTIONS

(a) 4 amps. If the original current was 1 amp and the number of electrons increased by four times the new current must be 4 amps.

(b) 3 amps × 1.4 = 4.2 amps.

(c) 7 amps × 1.4 = 9.8 amps.

(d) 7 amps ÷ 1.4 = 5 amps.

(e) 21 amps ÷ 1.4 = 15 amps.

(f) $6.00 = 600 cents.

(g) 350 cents = $3.50.

(h) 2 amps = 2000 ma. To perform the conversion we write 2 amps as 2.000 and then move the decimal point three places to the right and get 2000. In performing a conversion of this type, another way to look at it simply is adding three zeros.

(i) 6 amps = 6000 ma.

(j) 3.5 amps = 3500 ma. Again, we have simply moved the decimal point three places to the right; 3.5 amps = 3500 ma.

(k) .42 amps = 420 ma.

(l) .037 amps = 37 ma. Moving the decimal point three places to the right, we get 037, and since the zero has no significance we simply write the answer as 37 ma.

(m) .002 amps = 2 milliamps. Moving the decimal point three places to the right we get 002. ma and we drop the two zeros to the left of the 2 since they have no significance.

(n) 46 ma = .046 amps. Here we move the decimal point three places to the left. In order to do this we add a 0 to the left of the 4; thus 46 ma becomes .046 amps.

(o) 822 ma = .822 amps. To convert from the smaller unit to the larger unit we move the decimal point three places to the left.

(p) 1327 ma = 1.327 amps. Again, we simply move the decimal point three places to the left.

(q) 2 amps = 2,000,000 μa. To convert from amps to μa we have moved the decimal point six places to the right. To do this we must add the six zeros.

(r) .0017 amps = 1700 μa. Moving the decimal point six places to the right, we have to add two zeros to the right of the 7 in order to do this and .0017 can be written .001700 and then we move the decimal point six places to the right and get 1700 μa.

(s) 20 μa = .00002 amps. To convert from the smaller unit to the larger unit we move the decimal point six places to the left. In order to move it six places we have to add four zeros to the left of the 2.

(t) 147 μa = .000147 amps. Again, we add zeros to the left of the 1 and move the decimal point six places to the left to convert from microamps to amps.

(u) .26 ma = 260 μa. To convert from milliamps to microamps we move the decimal point three places to the right. In order to move it three places we have to add a zero to the right of the 6 so .26 ma becomes 260 μa.

(v) .031 ma = 31 μa. Again, to convert from the larger unit ma to the smaller unit microamperes we move the decimal point to the right. Thus, .031 ma becomes 031. μa. The 0 to the left of the 3 has no significance so we drop it and write the answer as 31 μa.

(w) 6100 μa = 6.1 ma. To convert from the smaller unit to the larger unit we move the decimal point to the left - thus when we move it three places to the left 6100 μa becomes 6.1 ma.

(x) 927 μa = .927 ma. Moving the decimal point three places to the left 927. μa becomes .927 ma.

(y) 327,000 μa = 327 ma. To convert from the smaller unit to the larger unit we move the decimal point three places to the left. Thus, 327,000 μa becomes 327 ma. To convert the 327 ma to amps we again move the decimal point three places to the left and get .327 amps.

(z) - 45 volts.

(aa) Terminal 1 is + 45V, and terminal 2 is - 45V.

(ab) The peak voltage between terminal 1 and ground will be 20 × 1.4 = 28 volts. It will be positive during one half cycle and negative the next half cycle.

(ac) + 60 volts. The two batteries are connected in series aiding, and therefore their potentials add, 15 + 45 = 60 volts. The terminal is positive because it's connected directly to the positive terminal of one of the batteries.

(ad) 0 volts. The two batteries are connected in series opposing and therefore the battery voltages subtract. Since both batteries are 22-1/2 volts the net value of the voltage between terminal 1 and ground will be 0.

(ae) + 22-1/2 volts. The two batteries are connected to oppose each other; thus their voltages subtract. 45V - 22-1/2V = 22-1/2V. The polarity will be that of the higher voltage battery, which is the 45 volt battery. Since its positive terminal is connected to terminal 1, then terminal 1 will be positive.

(af) - 1.5 volts. The two batteries are connected in series opposing and therefore their voltages subtract, 4.5 volts - 3 volts = 1.5 volts. The polarity will be that of the larger battery and since the negative terminal of the 4.5 volt battery connects to terminal 1, terminal 1 will be negative. Reversing the batteries had no effect on the polarity because we simply reversed their positions, keeping their polarities as shown in Fig. 9C. It makes no difference which position the battery is in insofar as the total voltage is concerned; as long as they are opposing each other you subtract to get the voltage produced by the two in series. In the circuit as shown in Fig. 9C, the polarity of terminal 1 is - 1.5 volts just as it is when the position of the two batteries is changed.

(ag) 30 volts when the two are aiding and 0 volts when the two are opposing. When the peak generator voltage has the same polarity as the battery voltage the total voltage will be the sum of the two voltages, 15V + 15V = 30V. When the generator has the opposite polarity to the battery, the two voltages will subtract. 15V - 15V = 0V.

(ah) +35 volts when they are aiding and - 5 volts when they are opposing.

When the peak generator voltage has the same polarity as the battery voltage, the two voltages add, 20V + 15V = 35V. When the two voltages oppose, the polarity of the generator will be - 20V with respect to terminal 1. Therefore the voltage between terminal 1 and ground will be 15V - 20V = - 5V.

(ai) - 75V when they aid and + 15V when they oppose. When the peak generator voltage has the same polarity as the battery voltage, the two voltages add. 30V + 45V = 75V. Since the negative terminal of the battery is connected to terminal 1, it will be - 75V with respect to ground. When the generator voltage opposes the battery voltage the peak voltage between terminal 1 and ground will be - 30V + 45V = + 15V.

(aj) + 30 volts when they are aiding and + 10 volts when they are opposing. When the peak generator voltage aids the battery voltage, the two voltages add. + 20V + 10V = + 30V. When the peak generator voltage opposes the battery voltage will have, + 20V - 10V = + 10V.

(ak) - 75V when they are aiding and

- 15V when they are opposing. When the two voltages have the same polarity their voltages add. 45 + 30 = 75V and since the negative terminal of the battery connects to terminal 1, terminal 1 will be negative. Another way of looking at this is to write the battery voltage as - 45V and the generator voltage as - 30V and add the two together. - 45V - 30V = - 75V. When the two voltages oppose the generator voltage subtracts from the battery voltage. 45V - 30V = 15V. Since the battery voltage is higher than the generator voltage the polarity of terminal 1 will be the polarity of the battery which is minus and therefore it will be -15V. Another way of doing this is to write the voltages down with their polarity. Here we have -45V + 30V = - 15V.

(al) The current will decrease. As a matter of fact, if we double the resistance in the circuit, the current will be cut exactly in half.

(am) Carbon resistors, wire-wound resistors and metal oxide film resistors are the three most widely used types of resistors in electronics.

(an) 4.7 K-ohms. To convert 4,700 ohms to K-ohms, you move the decimal point three places to the left. This is the same as dividing by 1,000 and getting 4.7 K-ohms.

(ao) 5.6 megohms. To convert 5,600,000 ohms to megohms you must move the decimal point six places to the left. This is the same as dividing by 1,000,000.

(ap) 330K-ohms. There are 1,000 ohms in a K-ohm and 1,000,000 ohms in a megohm. Therefore there must be 1,000 K-ohms in a megohm. To convert megohms to K-ohms you multiply by 1,000 and to do this you move the decimal point three places to the right. Thus, .330M = 330K.

(aq) 2,200,000 ohms. There are 1,000,000 ohms in a megohm and therefore to convert 2.2 megs to ohms you must multiply by 1,000,000. To do this you move the decimal point six places to the right.

(ar) 8,200 ohms. There are 1,000 ohms in a K-ohm and therefore to convert K-ohms to ohms, you multiply by 1,000. You do this by moving the decimal point three places to the right.

(as) .680 megohms. There are 1,000 K-ohms in a megohm and therefore to convert K-ohms to megohms you must divide by 1,000. To do this you simply move the decimal point three places to the left.

(at)

The symbol for resistance is shown above. There are probably more resistors used in electronics equipment than any other parts, so it is extremely important that you remember this symbol. The resistance of a resistor is usually indicated by writing the resistance either directly above or directly below the resistance symbol. The value may be given in ohms, K-ohms or megohms depending upon the size of the resistor. Usually the shortest form is used in order to conserve space on the diagram.

(au) $R = \frac{E}{I}$

(av) $E = IR$

(aw) $I = \frac{E}{R}$

(ax) 5 amps. To solve this problem you use the formula:

$$I = \frac{E}{R}$$

and substituting 15 volts for E and 3 ohms for R we get

$$I = \frac{15}{3}$$

$$= 5 \text{ amps}$$

(ay) The current will be 2 ma. We use the formula:

$$I = \frac{E}{R}$$

Since dividing 6,000 into 12 would be a decimal division, we can multiply by 1,000 and get our answer directly in milliamperes. When we do this we have:

$$I \text{ (ma)} = \frac{12}{6{,}000} \times 1{,}000$$

$$I = \frac{12{,}000}{6{,}000}$$

and cancelling three zeros above and below the line we get:

$$I = \frac{12}{6}$$

$$= 2 \text{ ma}$$

(az) The current in this case will be 4 ma. You use exactly the same method as in the preceding example; since dividing 7,000 into 28 will involve a decimal division you can multiply by 1,000 and get your answer directly in milliamperes. In this problem we have:

$$I = \frac{28}{7{,}000} \times 1{,}000$$

$$I = \frac{28{,}000}{7{,}000}$$

and then cancelling three zeros above and below the line we have:

$$I = \frac{28}{7}$$

$$= 4 \text{ ma}$$

(ba) In this problem if we multiplied by 1,000 to get our answer in milliamperes, we would still have a decimal division because we must convert 300K to ohms by multiplying it by 1,000. Thus we would have:

$$I = \frac{15}{300 \times 1000} \times 1{,}000$$

The 1,000 above the division line would simply cancel the 1,000 below the division line and we would have to divide 300 into 15. So instead of converting our answer directly to milliamperes it would be better to convert to microamperes. Now the problem becomes:

$$I = \frac{15}{300 \times 1000} \times 1{,}000{,}000$$

$$I = \frac{15,000,000}{300,000}$$

and now cancelling five zeros above the line and five zeros below the line we have

$$I = \frac{150}{3} = 50\ \mu a$$

(bb) The voltage will be 18 volts. We use the formula:

$$E = IR$$

and substituting 3 amps for I and 6 ohms for R we get:

$$E = 3 \times 6$$
$$= 18 \text{ volts}$$

(bc) In this example the current is 20 ma, and we must convert this to amps. The easiest way to do this is to simply divide it by 1,000. Therefore we will substitute $\frac{20}{1000}$ for I in the formula and 1000 for R. Using the formula:

$$E = IR$$

$$E = \frac{20}{1000} \times 1,000$$

$$E = \frac{20,000}{1,000}$$

and now we simply cancel three zeros above the line and three zeros below the line and we get:

$$E = 20 \text{ volts}$$

(bd) In this problem, the current which is in milliamps must be converted to amps by dividing it by 1,000 and the resistance which is in K-ohms must be converted to ohms by multiplying it by 1,000. Using the formula to find the voltage we have:

$$E = \frac{18}{1000} \times 3 \times 1000$$

This can be written:

$$E = \frac{18 \times 3 \times 1000}{1000}$$

and now we simply cancel the 1000 above the line and the 1000 below the line and we get as our answer:

$$E = 18 \times 3$$
$$= 54 \text{ volts}$$

(be) In this example the current which is given in microamps must be converted to amps by dividing it by 1,000,000. Thus our problem becomes:

$$E = \frac{47}{1,000,000} \times 200,000$$

Now you can cancel five zeros above the line and five zeros below the line and we get:

$$E = \frac{47 \times 2}{10} = \frac{94}{10} = 9.4 \text{ volts}$$

(bf) In this example, the current which is in microamperes must be converted to amps by dividing it by 1,000,000, and the resistance which is given in K-ohms must be converted to ohms by multiplying it by 1,000. Thus using the formula to find the voltage we have:

$$E = \frac{58}{1,000,000} \times 330 \times 1,000$$

$$E = \frac{58 \times 330 \times 1000}{1,000,000}$$

We have four zeros above the division line so we can cancel four zeros below the line and then our problem becomes:

$$E = \frac{58 \times 33}{100}$$

and now multiplying 58 × 33 we get:

$$E = \frac{1914}{100} = 19.14 \text{ volts}$$

(bg) In this example the current, which is in μa, must be converted to amps by dividing it by 1,000,000, and the resistance which is in megohms, must be converted to ohms by multiplying it by 1,000,000. Thus to find the voltage we have:

$$E = \frac{6}{1,000,000} \times 2 \times 1,000,000$$

$$E = \frac{6 \times 2 \times 1,000,000}{1,000,000}$$

and now we can simply cancel the 1,000,000 above the division line with the 1,000,000 below the division line and our voltage becomes

$$E = 6 \times 2 = 12 \text{ volts}$$

(bh) 3 ohms. To solve a problem of this type we use the formula:

$$R = \frac{E}{I}$$

In substituting 12 volts for E and 4 amps for I we get:

$$R = \frac{12}{4}$$

$$= 3 \text{ ohms}$$

(bi) In this example, the current is given in milliamperes and we must convert it to amps. We can do this by multiplying the voltage by 1,000. When we do this we have:

$$R = \frac{24 \times 1,000}{8}$$

$$R = \frac{24,000}{8}$$

$$= 3,000 \text{ ohms}$$

(bj) We use the same procedure in this example as we did in the preceding example. We multiply the voltage by 1,000 and this is the equivalent of converting the current which is in ma to amps. Therefore our problem becomes:

$$R = \frac{45 \times 1,000}{15}$$

$$R = \frac{45,000}{15}$$

$$= 3,000 \text{ ohms}$$

(bk) In this example, the current is given in microamperes. To convert the current to amperes we can multiply the voltage by 1,000,000 and then proceed with our problem.

$$R = \frac{34 \times 1,000,000}{170}$$

$$R = \frac{34,000,000}{170}$$

$$R = \frac{3,400,000}{17}$$

$$= 200,000 \text{ ohms.}$$

(bl) This example is the same as the preceding example; the current is in microamperes and we must convert it to amps. We do this by multiplying the voltage by 1,000,000. Our problem therefore becomes:

$$R = \frac{144 \times 1,000,000}{120}$$

$$R = \frac{144,000,000}{120}$$

$$R = \frac{14,400,000}{12}$$

$$R = 1,200,000 \text{ ohms.}$$

If we wish to convert this value to megohms we can do so by dividing by 1,000,000, in which case our answer becomes 1.2 megs.

(bm) The completed statement, which is very important, is as follows: If the voltage applied to a circuit is increased, the current will increase, and if the voltage applied to a circuit is decreased, the current will decrease.

(bn) The complete statement, which describes how resistance affects the current in a circuit, is as follows: if the resistance in a circuit is increased, the current will decrease, and if the resistance in a circuit is decreased, the current will increase.

LOOKING AHEAD

You have reached a point where you have learned about voltage, current and resistance and have started to study simple circuits. You used the three forms of Ohm's Law in performing calculations in simple circuits.

In the next lesson, you will put some of these simple circuits together to form more complex circuits and you will see that even more complex circuits follow the same rules that were followed in these simple circuits. You have already come a long way in your studying of electronics and you will add considerably to your knowledge in the next lesson.

Lesson Questions

Be sure to number your Answer Sheet B103.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

15 1. If a 9 volt battery and a 6 volt battery are connected in series aiding, what will the output voltage of the two batteries be?

3 2. If a 6 volt battery and a 3 volt battery are connected in series opposing, what will the output voltage be?

3. If a battery and a generator are connected as shown in Fig. 10A, and the battery voltage is 6 volts and the peak generator voltage is 6 volts, what will the maximum positive potential reached at terminal 1 be?

4. If in the circuit shown in Fig. 10A, the battery voltage is 6 volts and the generator peak voltage is 9 volts, will terminal 1 ever be negative with respect to ground - if so how much?

5. What are the three types of resistors commonly used in electronic equipment?

6. Give the three forms of Ohm's Law.

7. If the current flowing in a circuit is 2 amps, when the voltage is 50 volts, what will the current be if the voltage is reduced to 25 volts?

8. If the current flowing in a circuit is 50 ma when the resistance is 1,000 ohms, what will the current be if the resistance is reduced to 500 ohms?

9. If the voltage applied to a circuit is 4 volts, and the resistance in the circuit is 2,000 ohms, what will the current be?

10. If the current flowing in a circuit is 3 ma and the resistance is 6,000 ohms, what will the voltage be?

GETTING YOUR 'SECOND WIND'

After an opening burst of speed, a champion long-distance runner drops down to a steady natural pace. This brief period of relaxation releases that reserve of power called "second wind." He is then able to overtake and pass the now nearly exhausted leaders in the race.

No matter what you are doing, you can always do it better once you get your second wind. Instead of fighting that sleepy feeling which sometimes comes when you are studying, stop and relax for a few minutes so as to release your own supply of reserve power. Here are some ways to relax, all worth trying.

Get up and exercise for a few minutes. Get a drink of cold water. Step outside for a few deep breaths of fresh air. Take a brisk walk up and down the road or once around the block.

If you "shake yourself awake" in one of these ways, you'll find it isn't at all hard to get that second wind which enables you to study longer and makes the going easier. Then you'll get in some real worth-while studying -- then you'll get things done!

J. C. Thompson

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-203

ACHIEVEMENT THROUGH ELECTRONICS

SERIES AND PARALLEL CIRCUITS

B104-1

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

SERIES AND PARALLEL CIRCUITS

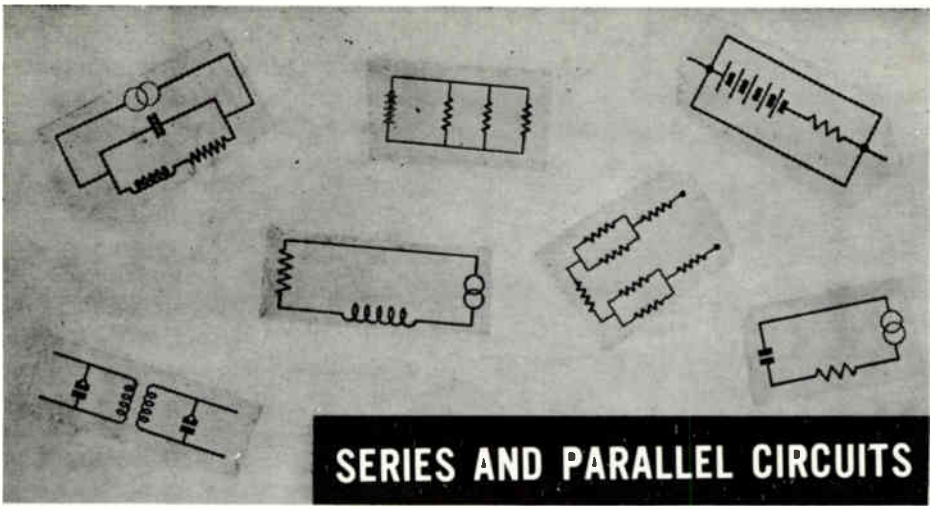
B104-1

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

NATIONAL RADIO INSTITUTE, WASHINGTON, D.C. 20016

P115M575 | 1975 EDITION | Litho in U.S.A.



So far you have been studying relatively simple circuits. However, in spite of this, you have learned a great deal about electricity. You have studied voltage, current and resistance and the units that are used to measure them. You know that the units in which current is measured are the ampere, milliampere and microampere. The units used to measure voltage are the volt, the millivolt, the microvolt and the kilovolt. The units used to measure resistance are the ohm, K-ohm and megohm. You should be familiar with all these quantities at this time; you should know what they mean and how to convert from one unit to another.

You have also studied Ohm's Law and how it relates voltage, current and resistance. If you know any two of these values, you should be able to use Ohm's Law to find the third.

So far the circuits that you have studied have been relatively simple circuits. In almost all cases they have been series circuits. By a series circuit, we mean that the parts are connected so that there is only one path through which current can flow. In other words, if a resistor is connected between the negative and positive terminals of a battery, the electrons must leave the negative terminal of the battery, flow through the resistance to the positive terminal of the battery and then through the battery back to the negative terminal. There is only one path through which the current can flow.

There are other series circuits in addition to a circuit with only one resistor. A series circuit may have a number of resistors such as the circuit shown in Fig. 1. The circuit is called a series circuit, because like the simple circuits you have studied, the parts are connected in series so that electrons must flow first from the negative terminal of the battery through R_1, and then through R_2 and finally through R_3 back to the positive terminal of the battery and then through the battery back to the negative terminal. Electrons flow through one part after the other, in series. The series circuit is a very important circuit in elec-

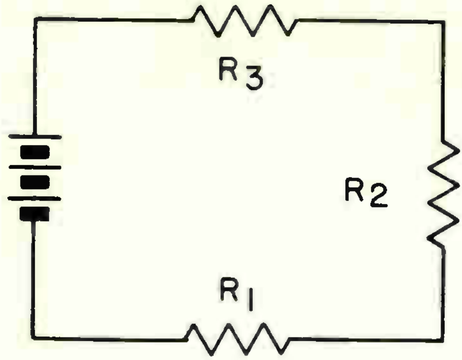


Fig. 1. A series circuit.

tronics; make sure that you understand what is meant by a series circuit.

In addition to a series circuit, we have what is known as a parallel circuit, such as is shown in Fig. 2. Here we have three resistors connected in parallel across the battery. We say that they are in parallel because there are three parallel paths through which current from the battery can flow. Electrons leaving the negative terminal of the battery can flow through either R_1, R_2, or R_3 and then back to the positive terminal of the battery and through the battery back to the negative terminal.

In this lesson you are going to study both series and parallel circuits. Both are important in electronics. Many circuits found in electronic equipment are series circuits; many circuits are parallel circuits. Some circuits are combinations of both series and parallel circuits.

When resistors are connected in series, such as in Fig. 1, you sometimes need to find the total resistance of the resistors. Similarly when they are connected in parallel as in Fig. 2, sometimes you have to find the effective value of the three resistors in parallel. We will learn how to do both in this lesson and you will learn how current flow is affected by connecting resistors both in series and in parallel.

In the first section of the lesson we will go into series circuits in considerable detail, and then in the second section we will go into parallel circuits. After we have completed both series and parallel circuits, we will study circuits that are combinations of both series and parallel circuits. As you may well realize, it is extremely important that you understand everything covered in series circuits first, and then what is covered in parallel circuits later. If there are some points that give you trouble, be sure to restudy the parts of the lesson you are in doubt about before going on. If you cannot work out the problem yourself, be sure to write to NRI and get additional help. If you do not master series circuits you will have trouble with parallel circuits. If you leave some points unclear in parallel circuits, then when you go on to study combinations of series and parallel circuits you will have difficulty with

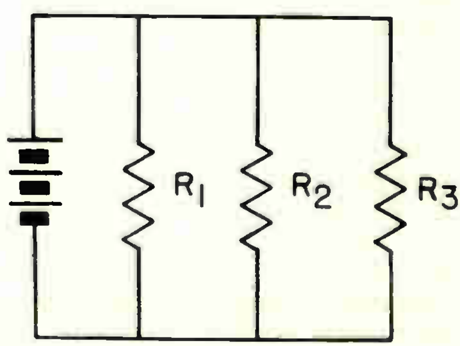


Fig. 2. A parallel circuit.

them. Learning electronics is simply a matter of taking one simple circuit at a time, learning how it works and then building on that knowledge. If you miss some of the basic fundamentals, you will have difficulty with more advanced circuits.

Be sure you can answer the self-test questions at the end of each section. These questions will test you on each section, and they are a good indication of whether or not you have mastered the material covered. Now, let us go ahead and study series circuits.

Series Circuits

A series circuit is a circuit in which there is only one path for electrons to flow. In a series circuit, such as we saw in Fig. 1, the current flowing is the same at all parts in the circuit. In other words, the current flowing through R_1 is equal to the current flowing through R_2; equal to the current flowing through R_3 and equal to the current supplied by the battery. The leads or conductors connecting the resistors and batteries together are also carrying the same current as is flowing through the resistors. This is one of the important things you should remember about a series circuit. There is only one path for electrons to flow, and the current flow in the circuit must be the same at all points in the circuit.

When we have resistors connected in a circuit such as in Fig. 1, we say that the resistors are connected in series. Now, let us see what effect these resistors connected in series have on the current flow in the circuit.

RESISTORS IN SERIES

In Fig. 3 we have shown a circuit that is identical to Fig. 1, except that in Fig. 3 we have given the battery a voltage of 15 volts and given each resistor a definite value. When the resistors are connected in series in this way, each resistor opposes the flow of current in the circuit. The total opposition to current flow will be equal to the sum of the resistances of the individual resistors. Thus, when resistors are connected in series, we can say that their total resistance is equal to the sum of the resistances. In this particular case we can express this mathematically as:

$$R_T = R_1 + R_2 + R_3$$

This formula can be used for find-

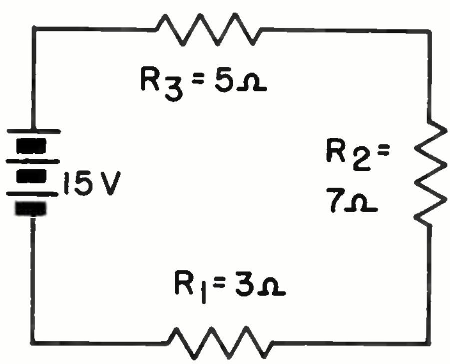


Fig. 3. A series circuit.

ing the resistance of any number of resistors connected in series. The total resistance is always equal to the sum of the individual resistors. Thus if you have five resistors connected in series and want to know the total resistance of the five in series you simply add the resistances together. Similarly, if there are ten resistors connected in series, to find the total opposition to current flow you simply add the value of the ten resistors together.

The opposition to current flow in the circuit will be equal to the total resistance of the resistors in series. In Fig. 3, the total resistance is:

$$R_T = 3 + 7 + 5 = 15 \text{ ohms}$$

Therefore, the three resistors connected in series offer the same opposition to current flow in the circuit as a single 15-ohm resistor would offer.

To find the current that will flow in the circuit you use Ohm's Law. You use it in the form:

$$I = \frac{E}{R}$$

When you have a number of resistors in series, R becomes the total value of these resistors, or R_T. In the circuit shown in Fig. 3, $R_T = 15$ ohms, and therefore in this circuit, the total current flow will be:

$$I = \frac{15}{15} = 1 \text{ amp}$$

The example given in Fig. 3 was comparatively simple, because the resistance values were all small and we did not get involved in any decimal divisions in finding the current. A somewhat more difficult example of a series circuit is shown in Fig. 4. However, this should not cause any more difficulty in finding the total resistance in the circuit than the example in Fig. 3, nor should it cause you any more difficulty in finding the total current flowing in the circuit, if you use the procedures you learned in the preceding lessons.

The first step in finding the total

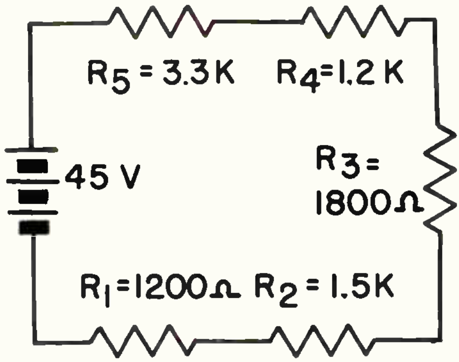


Fig. 4. A series circuit with five resistors.

current that will flow in this circuit is to find the total resistance. To do this, you must convert all the resistance values to ohms. $R_1 = 1200$ ohms. $R_2 = 1.5$K-ohms. To convert this to ohms you multiply by 1000; thus $1.5 \times 1000 = 1500$ ohms. R_3 is given as 1800 ohms. R_4 is 1.2K-ohms, and again to convert this to ohms you multiply by 1000 and get 1200 ohms. Similarly, R_5 which is 3.3K-ohms is equal to 3300 ohms.

Now you simply write down the values of these resistances as shown in Fig. 5, and then add the five values to get the total resistance in the circuit.

As shown in Fig. 5, the total resistance of the five resistors connected in series is 9000 ohms. Now, we use Ohm's Law to find the current that will flow in the circuit. We use the form,

$$I = \frac{E}{R}$$

If we substitute the values of 45 volts for the voltage, and 9000 ohms for the resistance, we will see that we immediately are involved with a decimal division. However, you will remember from the preceding lesson that you could multiply by 1000 and get the current directly in milliamperes. Once you do this the problem becomes:

$$I = \frac{45}{9000} \times 1000$$

and we can cancel three zeros above and below the division line so that our problem is

$$I = \frac{45}{9}$$

$$= 5 \text{ ma}$$

Any series circuit problem can be solved in the same way. To find the current flowing in the circuit you find the total resistance in the circuit and then divide this total resistance into the voltage applied. This will give you the total current flow in the circuit.

MEASURING VOLTAGES

Voltages are measured by means of an instrument called a voltmeter. To measure the voltage across a battery, a generator or a circuit, you connect the voltmeter directly across the part where you want to measure the voltage.

$$\begin{array}{r} 1200 \\ 1500 \\ 1800 \\ 1200 \\ \underline{3300} \\ 9000 \end{array}$$

Fig. 5. Total value of the five resistors in Fig. 4.

To measure dc voltages, you use a dc voltmeter. A dc voltmeter has one terminal marked with a minus sign and the other terminal marked with a plus sign. To measure the voltage across the battery shown in Fig. 6, you connect the minus terminal of the meter to the negative terminal of the battery and the plus terminal of the meter to the positive terminal of the battery. When you connect a meter across a 30-volt battery such as shown in Fig. 6, the meter pointer will move up the scale and indicate a voltage of 30 volts. Don't worry about how the meter works now - we will go into that later. The important point to see here is that you connect the minus terminal of the meter to the minus terminal of the battery and the plus terminal of the meter to the plus terminal of the battery.

Now looking at Fig. 6, you see that

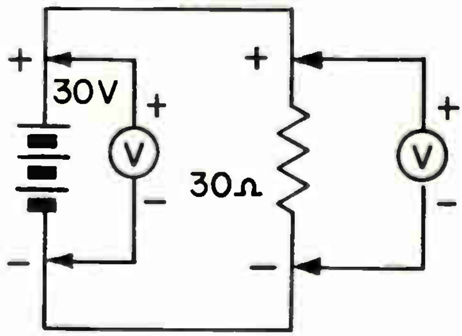


Fig. 6. Measuring voltage in a simple series circuit.

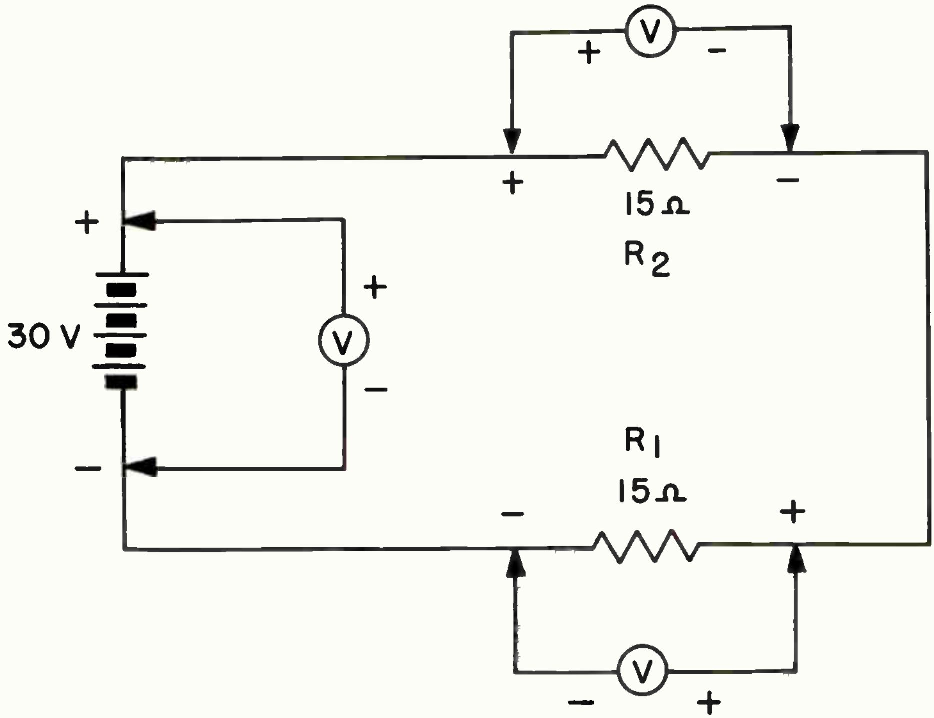


Fig. 7. Measuring voltage across series resistors.

the resistor is connected directly across the battery. Suppose we wanted to measure the voltage across the resistor, how would we do it?

Since the resistor is connected directly across the battery we would actually be reading the battery voltage if we connect the meter across the resistor. Therefore the meter must be connected across the resistor as shown in Fig. 6. The minus terminal of the meter must be connected to the end of the resistor that connects to the negative terminal of the battery and the plus terminal of the meter must be connected to the end of the resistor that connects to the plus terminal of the battery. Again, since the meter and resistor are both connected directly across the battery, the meter would indicate that the voltage was 30 volts.

Now, let us look at Fig. 7. Here, instead of a single 30-ohm resistor, we have two 15-ohm resistors connected in series across the 30-volt battery. Since the total resistance offered by the two 15-ohm resistors will be 30 ohms, insofar as the battery is concerned, the circuit will appear exactly like the one in Fig. 6. If we use Ohm's Law we would find that the current in both cases was 1 amp. If we measure the voltage across the battery, we connect the meter across the battery in exactly the same way as we connected it in Fig. 6.

If we want to measure the voltage across the two resistors in series, we would connect the negative terminal of the meter to the terminal of R_1 that connects to the negative terminal of the battery and the plus terminal of the meter to the terminal of R_2 that connects to the plus terminal of the battery. The meter would indicate 30 volts across the two resistors, because once again we are in effect simply connecting the meter across the battery. However, if the voltage across the two resistors in series is 30 volts, then it is logical to assume that the voltage across one resistor is only half this value or 15 volts. Indeed, if we connect the meter across R_1 we would find that the voltage is 15 volts, and if we can connect the meter across R_2 we would find that the voltage across R_2 is 15 volts.

To measure the voltage across one of these resistors, we connect the meter as shown. Notice that when measuring the voltage across R_1, the minus terminal of the meter must be connected to the end of the resistor that connects to the negative terminal of the battery, and the plus terminal of the meter is connected to the other end. This indicates that there is a voltage across R_1 having a polarity as shown in Fig. 7. To measure the voltage across R_2, we connect the meter with the plus terminal to the end of the resistor that connects to the positive terminal of the battery and the minus terminal of the meter to the other end of R_2. This indicates that the voltage across R_2 has the polarity shown.

We know that the current flow in this circuit will be 1 amp. We can prove that there should be a voltage of 15 volts across each of these resistors by using Ohm's Law. Ohm's Law states that if you know the resistance and current in a circuit you can find the voltage. You know that the resistance of R_1 is 15 ohms and you know that the current flowing through it is 1 amp. Putting these values in Ohm's Law we get:

$$E = IR$$
$$E = 1 \times 15 = 15 \text{ volts}$$

In the same way you could calculate the voltage across R_2 and you would find that it is also 15 volts. However, it is not necessary to do this because you know that the voltage across the series combination of R_1 and R_2 must be 30 volts since they are connected across the 30-volt battery, and that if the voltage across one of the resistors is 15 volts, the voltage across the other must be 15 volts also. In addition, in any circuit where you had equal resistors and the same current flows through the resistors, the voltage drops across the resistors must be equal.

We have taken this example one step further in Fig. 8. Here we have put three 10-ohm resistors in series across the 30-volt battery. Once again, you know that the total resistance in the circuit will be the sum of the individual resistors, which again is 30 ohms. Therefore the current flow in the circuit will be 1 amp and the voltage across each resistor will be 10 volts. To measure these voltages, you connect the meter across the individual resistors as shown. Notice that the negative or minus terminal of the meter always connects to the end of the resistor closest to the negative terminal of the battery and the positive or plus terminal of the meter connects to the end of the resistor closest to the positive terminal of the battery.

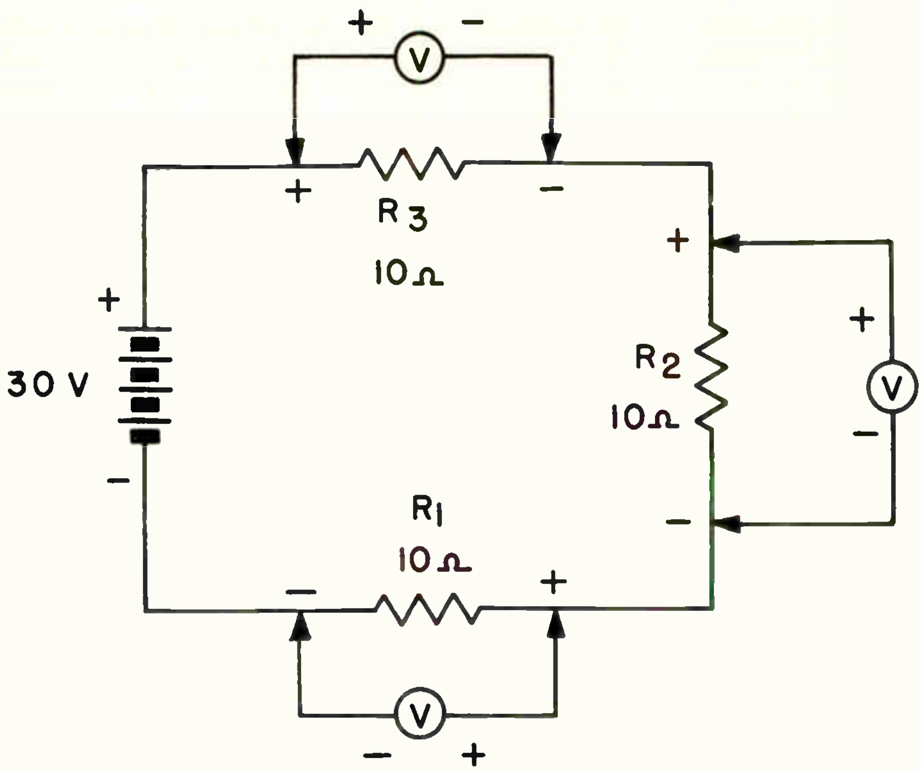


Fig. 8. Voltage polarities across three series resistors are as shown.

Across each resistor there will be a voltage, and in this case the voltage across each resistor will be 10 volts. The voltage across each resistor will have the polarity shown on the diagram.

VOLTAGE DROP

In referring to the voltage across the resistors in a circuit such as Fig. 8, we say that part of the voltage is used or dropped across each resistor. We refer to the voltage across a resistor as a "voltage drop". It is important that you remember this term, since we will use it over and over again. We say that the voltage is "dropped" across a resistor, and we refer to the voltage across each resistor as a "voltage drop".

In a series circuit such as the one shown in Fig. 8, the sum of the voltage drops will always be equal to the source voltage. This means that the voltage drop across R_1, plus the voltage drop across R_2, plus the voltage drop across R_3 will be equal to the battery voltage. This is true regardless of what the battery voltage is and regardless of the value of the resistors used in the series circuit. The sum of the voltage drops in a series circuit will always be equal to the source voltage.

Sometimes we have occasion to trace through a series circuit such as the one shown in Fig. 8 and record the voltages across the individual parts in the circuit. Let us do this starting at the minus terminal of the battery. The first part we come to is R_1 and the voltage across R_1 is 10 volts. Now the question becomes, should we write this as minus 10 volts or plus 10 volts? In tracing through the series circuit, we trace from the minus end of R_1 to the plus end. Let us say that all voltage encountered tracing from minus to plus will be indicated as positive voltages. Therefore the voltage drop across R_1 is + 10 volts. Now we come to R_2, and the voltage across it is also 10 volts. Since we are tracing from the minus end of the resistor to the plus end we indicate the voltage across this resistor as + 10 volts too. Next, we come to R_3 and the voltage across it is 10 volts and we are tracing it in the same way, from the minus end of the resistor to the plus end so we indicate this voltage as + 10 volts. Now as we continue to trace through the circuit we come to the positive terminal of the battery. Now we trace through the battery from the positive terminal to the minus terminal - in other words we are tracing through the battery in the opposite direction from which we traced through the resistors. Therefore since we have called the other voltages plus voltages, we must indicate this as -30 volts. Now if we add the voltages around the circuit we have

+ 10 volts + 10 volts + 10 volts - 30 volts =

+ 30 volts - 30 volts = 0

This brings us to another important rule or law in electronics known as Kirchhoff's Law. It states that the sum of the voltages in a closed circuit is 0. In other words, when you trace around a complete series circuit such as is shown in Fig. 8, and add the voltages with the proper polarity, the sum of these voltages will be 0. This is essentially the same as the statement that the sum of the voltage drops in the circuit will be equal to the source voltage. We have gone through both statements even though they are essentially the same, because you will run into both in your studies of electronics and it's important that you know what they mean and know that they are really saying the same thing.

Knowing that the sum of the voltage drops in a series circuit is equal to the source voltage will enable you to find the voltage across a resistor in an example such as the one shown in Fig. 9. Here you are given the

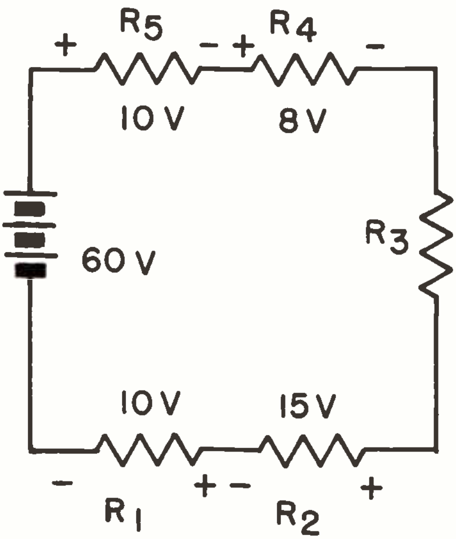


Fig. 9. The voltage across R_3 can be found in the above circuit.

source voltage as 60 volts. You are given the voltage across four of the resistors and want to find the voltage across R_3. You know that the sum of the voltage drops across the five resistors must be equal to 60 volts. So you add the voltage drops that are known across the four resistors. This will give you

$$10 + 15 + 8 + 10 = 43 \text{ volts}$$

Since the sum of the five voltage drops must be equal to 60 volts and the voltage drops across the four resistors total 43 volts, then you know that the voltage drop across R_3 must be equal to 17 volts.

Being able to work simple problems like this in series circuits will be helpful to you. In some electronic equipment parts may be buried in such a way that it is practically impossible to get at them with a voltmeter. You might want to measure the voltage across such a part. Sometimes, while it is impossible to get at the particular part across which you want to measure the voltage, you can measure the source voltage across this part and a number of other parts in series. Then if you can measure the voltage drop across the other parts and subtract these voltage drops from the source voltage, you can determine what the voltage drop is across the part in which you are interested. Being alert to such simple things as this is often what makes the difference between an expert technician who is able to get at the source of trouble in a piece of equipment quickly, and a technician who sort of blunders around trying first one thing and then the other, servicing more or less by a hit-and-miss procedure.

TUBE AND TRANSISTOR CIRCUITS

Often vacuum tube and transistor circuits are simple series circuits. In Fig. 10, we have shown two examples of such circuits. Looking at

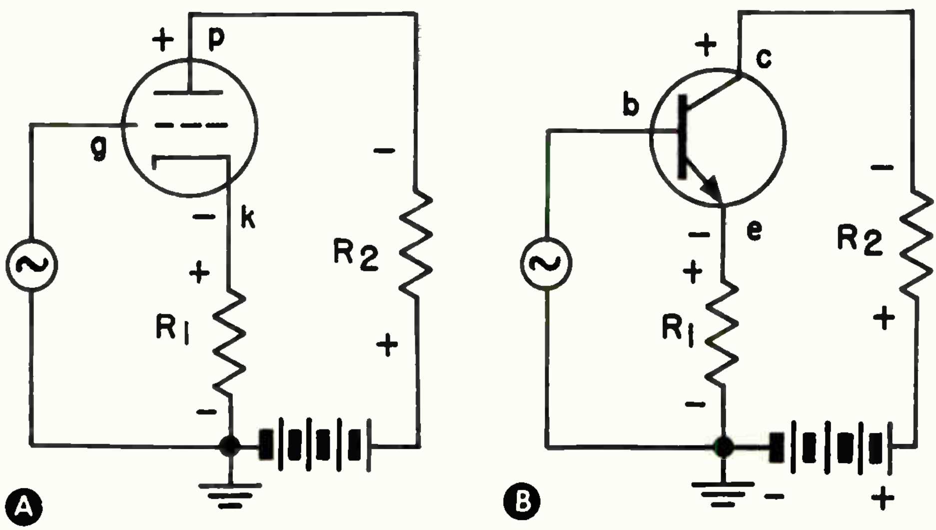


Fig. 10. A series tube circuit is shown at A; a series transistor circuit at B.

Fig. 10A first, here we see a triode vacuum tube. Some of the parts that are normally found in the circuit have been omitted for simplicity. The circuit that we are interested in is the series circuit made up of the battery, R_1, the tube and R_2. Electrons leave the negative terminal of the battery and flow through R_1 to the cathode of the triode tube. We use the letter k to designate the cathode. The cathode is heated to a red heat by the heater inside the cathode. (We have not shown the heater on the diagram because it is not part of the series circuit in which we are interested.) The electrons fly off the heated cathode and travel over to the plate of the tube. From the plate the electrons flow through R_2 back to the positive terminal of the battery and through the battery back to the negative terminal. Thus we have a series circuit consisting of R_1, the vacuum tube, R_2 and the battery. Across each of these parts there will be a voltage drop. The voltage drop across R_1 will have the polarity shown; the end that connects to the negative terminal of the battery will be negative and the end of the resistor that connects to the cathode of the tube will be positive. There will be a voltage drop between the cathode and the plate of the tube; the cathode will be negative and the plate positive. There will be a voltage drop across R_2, the end of the resistor that connects to the plate of the tube will be negative, and the end that connects to the positive side of the battery will be positive. By using a dc voltmeter and connecting it with the proper polarity we can actually measure these voltage drops, and if we use the right kind of meter so it will not upset the performance of the circuit, we will find that the sum of these voltage drops is equal to the battery voltage.

In Fig. 10B, we have shown a transistor circuit. We have also left out some of the parts in this circuit to simplify it. The series circuit we are interested in consists of R_1, the transistor, R_2, and the battery. In this circuit, electrons leave the negative terminal of the battery and flow through R_1 to the emitter (marked e) of the transistor. Here they flow from the emitter to the base (marked b), and from the base to the collector (marked c), and from the collector through R_2 back to the positive terminal of the battery, and through the battery back to the negative terminal. We have a series of voltage drops around the circuit as shown by the polarities marked on the diagram. Again, we can measure the voltage drop across R_1, the voltage drop between the emitter and collector of the transistor and the voltage drop across R_2, and the sum of these three voltage drops will be equal to the battery voltage.

These simple series circuits are typical of the series circuits you will find in all types of electronic equipment. You will find that the current flowing through these simple circuits is extremely important. The generator connected between the grid and ground of the tube circuit and between the base and ground in the transistor circuit causes the current through these circuits to vary, and this varying current is what makes it possible for the tube and the transistor to amplify the signal. The generator in each case is representing a small signal input, such as might be obtained from a microphone or from a preceding stage. An amplified signal will appear across R_2 in each case. We will go into detail on how these devices amplify later. The important thing for you to see at this time is the simple series circuits involved and to realize how important it is that you are able to recognize series circuits and remember the important facts about them.

SUMMARY

Series circuits are extremely important; you will find all kinds of them in electronic equipment. You must be able to recognize a series circuit. It is a circuit in which electrons leave the negative terminal of the voltage source and flow through a number of parts, one after the other, to the positive side of the voltage source and through the voltage source back to the negative terminal.

Remember what is meant by a voltage drop. It is the voltage that appears across any part in a series circuit. You must also be able to indicate the polarity of the voltage drop across a part. The end of the part that is closest to the negative terminal of the voltage source will be negative, and the end of the part that is closest to the positive terminal of the voltage source will be positive.

Remember the two important rules about the voltages in a series circuit. The sum of the voltage drops in a series circuit is equal to the source voltage. Another way of expressing the same rule is that if you add the voltages with the correct polarity, the sum of the voltages around a series circuit will be equal to 0.

If there are any parts of the preceding section that are not clear to you, it would be worthwhile to go over the entire section again, before trying the self-test questions. These early lessons are the most important lessons in your course. A thorough understanding of basic circuits will help you all the way through your course; it will make later lessons comparatively easy. On the other hand, if you do not understand these basic circuits, then you will have difficulty all the way through your course. After you have mastered the material covered in this section answer the self-test questions. If you find that you cannot answer one of the self-test questions, don't hesitate to go back and review. These questions are put in the lesson to help you determine whether or not you know as much as you should about the section. By going back and finding the answer to a self-test question that you cannot answer, you will be reviewing a part of the lesson that is not clear to you. This will help you master the important parts of the lesson.

SELF-TEST QUESTIONS

(a) What is a series circuit?
(b) How do you connect a dc voltmeter to a battery to measure the battery voltage?
(c) Draw a simple series circuit consisting of a battery and three resistors. Label the resistor that connects to the negative terminal of the battery R_1, and the resistor that connects to the positive terminal of the battery R_3. Label the resistor between R_1 and R_3, R_2.
(d) Indicate the polarity of the volt-

age drops across the resistors in the series circuit you have drawn.

(e) If R_1 equals 3 ohms, R_2 equals 4 ohms and R_3 equals 5 ohms, what is the total resistance of the three resistors connected in series?

(f) If the total current flowing in your series circuit is 2 amps, find the voltage drop across each of the three resistors in the circuit.

(g) What is the source voltage in the series circuit you have drawn?

(h) In a series circuit in which four resistors are connected across a battery, the battery voltage is 35 volts. The voltage drop across one resistor is 5 volts, across another resistor it is 7 volts, and across a third resistor it is 10 volts. What is the voltage drop across the fourth resistor?

(i) In a vacuum tube circuit such as shown in Fig. 10A, the voltage drop across R_1 is 3 volts, the voltage drop across R_2 is 125 volts, and the source voltage is 250 volts. What is the voltage between the plate and cathode of the tube?

(j) In a transistor circuit such as shown in Fig. 10B, the voltage drop across R_1 is .2 volts. The voltage drop across R_2 is 3.6. volts. The battery voltage is 6 volts. What is the voltage drop between the emitter and collector of the transistor?

(k) Draw a series circuit with a battery and five resistors like the one shown in Fig. 4. R_1 equals 120 ohms, R_2 equals 150 ohms, R_3 equals 180 ohms, R_4 equals 120 ohms and R_5 equals 330 ohms. The battery voltage is 18 volts. Find the total resistance in the circuit and then find the current flowing in the circuit, and finally find the voltage drop across each resistor and label the polarity of the voltage drop.

Parallel Circuits

You might think of a parallel circuit as the opposite of a series circuit. In a series circuit, current flows through one part after the other in the circuit. There is only one path for current flow and the current is the same in all parts of the circuit. In a parallel circuit, a different current can flow through each branch of the circuit. If two resistors are connected in parallel there are two paths through which current can flow; if three resistors are connected in parallel then there are three paths through which current can flow, and the current flowing in each path or branch of the circuit can be different.

Parallel circuits might seem a little confusing at first, but actually they are quite simple even though there are some important rules

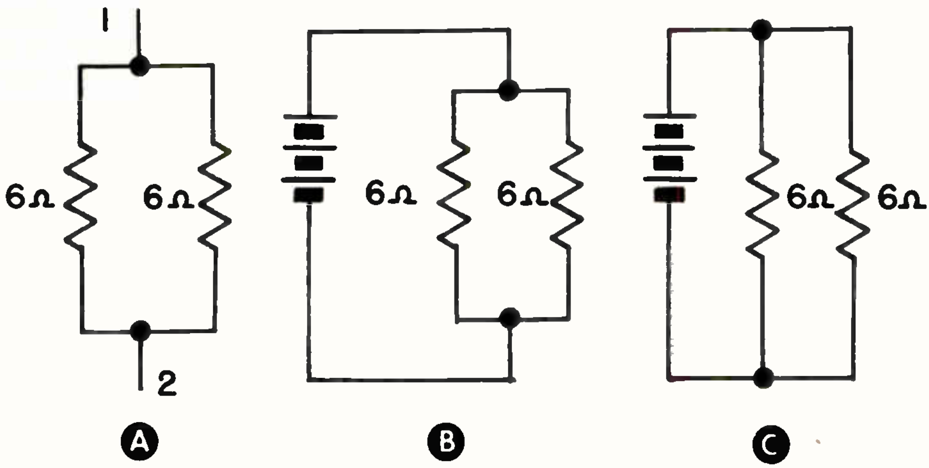


Fig. 11. Simple parallel resistor circuits.

which govern their performance. Once you learn these rules you will see that parallel circuits are no more difficult to understand than series circuits.

In the preceding section we learned that when two resistors were connected in series that their total resistance was equal to the sum of their resistances. We have a somewhat different situation in parallel circuits, so let us see exactly what happens.

RESISTORS IN PARALLEL

Often in electronic equipment two or more resistors may be connected in parallel. We need to be able to find the effective resistance of two resistors connected in parallel.

In Fig. 11A, we have shown two 6-ohm resistors connected in parallel. Between terminals 1 and 2 of these resistors there will be some net value of resistance. We can connect these two resistors in parallel across a battery as shown in Fig. 11B, and a certain current will flow in the circuit. The current that will flow in the circuit will depend upon the battery voltage, and the net resistance of the two resistors connected in parallel.

We can see better what is happening if we redraw the circuit as shown in Fig. 11C. Here we see clearly that each resistor is connected directly across the battery. Therefore there will be a current path from the negative terminal of the battery to the first 6-ohm resistor, through this resistor, and back to the positive terminal of the battery. The amount of current that will flow will depend upon the voltage of the battery, and the resistance of the resistor, which in this case is 6-ohms. There will be a similar path from the negative terminal of the battery and through the second resistor and back to the positive terminal of the battery.

Since the two resistors are of equal value, for a given battery voltage, equal currents will flow through the two resistors. Let us suppose that the battery voltage is 6 volts. Then, the current flowing through either of the 6-ohm resistors will be 1 amp. This means that the total current flowing in the circuit will be 2 amps, 1 amp through each resistor.

Now, since we know the battery voltage and the total current flowing, we can substitute these values in Ohm's Law and find the effective resistance in the circuit. Using

$$R = \frac{E}{I}$$

$$R = \frac{6}{2} = 3 \text{ ohms}$$

Thus the resistance of two 6-ohm resistors connected in parallel is 3 ohms - notice that it is one half the resistance of either resistor.

You can set up other examples of equal resistors connected in parallel, and you will find that it always works out that the total resistance of the resistors in parallel is equal to one half the resistance of either resistor.

You might wonder about our selecting a voltage of 6 volts to work out this problem. We selected 6 volts simply to make the division easy. We could select any voltage; the current through each resistor would be different, but the result will always be the same. For example, suppose we select a battery voltage of 18 volts. Then, from Ohm's Law, the current through each resistor will be three amps, and therefore the total current flowing in the circuit will be 6 amps. When we substitute these values into Ohm's Law we find that once again the total resistance in the circuit is 3 ohms. Selecting a voltage like this is an easy way of finding the resistance of two resistors connected in parallel: there is another method that can be used which we will show you later.

You can use the same procedure in finding the total resistance of two unequal resistors connected in parallel. For example, suppose we have a 4-ohm resistor and an 8-ohm resistor connected in parallel as shown in Fig. 12A. We can simply set up a simple circuit such as shown in Fig. 12B and then assume a convenient battery voltage which

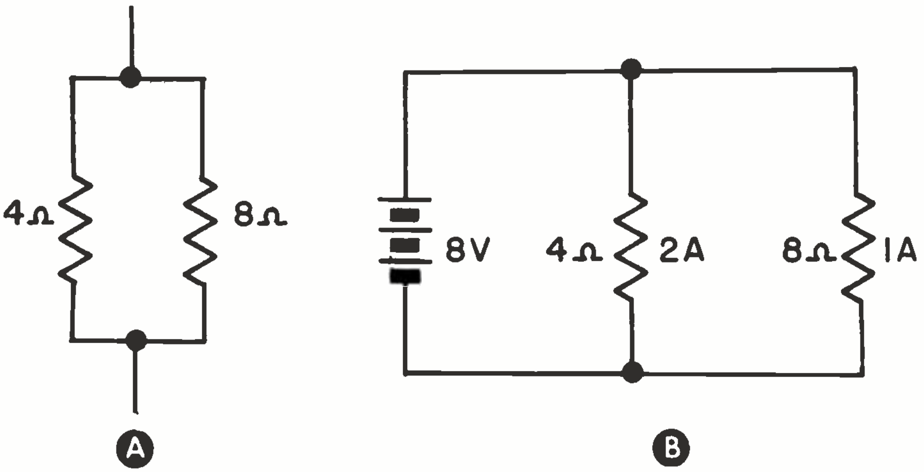


Fig. 12. Finding the value of 4Ω in parallel with 8Ω.

in this case might be 8 volts. With a voltage of 8 volts, a current of 2 amps will flow through the 4-ohm resistor and a current of 1 amp, will flow through the 8-ohm resistor. Therefore the total current flow in the circuit will be 3 amps. Now to find the total resistance we divide 3 amps into 8 volts and get 2.67 ohms.

In the examples shown in Figs. 11 and 12, it was easy to pick a voltage into which the resistance values can be divided to give convenient values of current. Sometimes this is not so easy and then we use another method of finding the value of resistors connected in parallel. We use the formula

$$R_t = \frac{R_1 \times R_2}{R_1 + R_2}$$

This formula states that the total resistance of two resistors connected in parallel is equal to the product of the resistance of the two resistors divided by the sum of the resistance of the two resistors.

For example, using the two 6-ohm resistors that we used in Fig. 11, in this formula we get

$$R_t = \frac{6 \times 6}{6 + 6}$$

$$R_t = \frac{36}{12} = 3 \text{ ohms}$$

This formula is particularly useful in finding the value of larger resistors. For example, find the value of 2400 ohms in parallel with 800 ohms. Substituting these values in the formula we get

$$R_t = \frac{2400 \times 800}{2400 + 800}$$

$$R_t = \frac{1920000}{3200}$$

and now cancelling two zeros above the division line and two zeros below the division line, we have

$$R_t = \frac{19200}{32} = 600 \text{ ohms}$$

Often you will have more than two resistors connected in parallel and want to find the total resistance of the group. In Fig. 13A we have shown three resistors in parallel. In Fig. 13B we have shown them connected across a battery. If we assume a battery voltage of 24 volts, then we will get a whole number for the current flowing through each resistor. The current through the 6-ohm resistor would be 4 amps, the current through the 8-ohm resistor would be 3 amps, and the current through the 12-ohm resistor would be 2 amps. This will give us a total current of 9 amps. Now using Ohm's Law we can find the total resistance

$$R = \frac{E}{I}$$

$$R = \frac{24}{9} = 2.666+ \text{ ohms}$$

which we round off to 2.7 ohms.

Another way to find the total resistance of three resistors in parallel is to use the formula first for two of the group, finding the parallel resistance, and then using the formula again with the result and the third resistor in parallel with the parallel resistance of the first two resistors.

In the example, if we find the resistance of the 6-ohm and 12-ohm resistor in parallel first, we will have less work to do with fractions. Using the formula, we find that 6 ohms in parallel with 12 ohms is

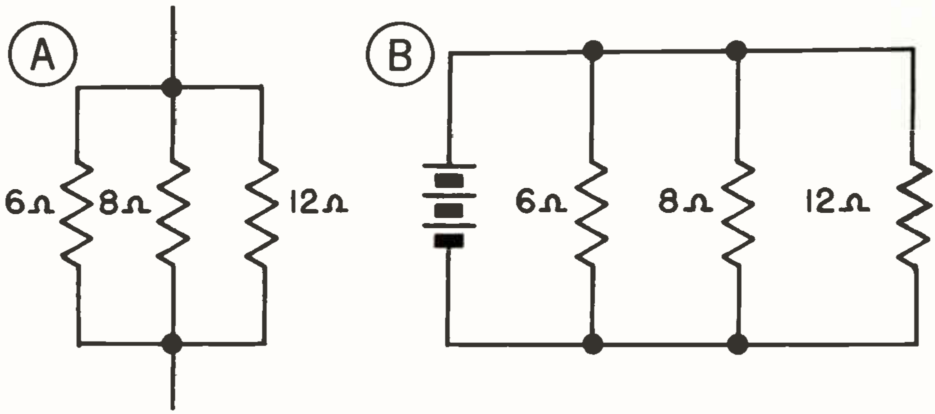


Fig. 13. A parallel circuit with three current paths.

$$R_t = \frac{6 \times 12}{6 + 12}$$

$$R_t = \frac{72}{18} = 4 \text{ ohms}$$

Thus the value of a 6-ohm resistor in parallel with a 12-ohm resistor is 4 ohms. Now we find the value of 4 ohms in parallel with 8 ohms, which we did previously and we find

$$R_t = \frac{4 \times 8}{4 + 8}$$

$$R_t = \frac{32}{12} = 2.666+ \text{ ohms}$$

Again we round it to 2.7 ohms.

If you have four or five resistors connected in parallel you can use the same formula, simply grouping them together in pairs and then substituting the parallel resistance in the formula. For example, in the case of five resistors you might find the value of R_1 and R_2 in parallel and then the value of R_3 and R_4 in parallel. Next, find the parallel resistance of the R_1-R_2 combination in parallel with the R_3-R_4 combination, and when you get the answer to the value of these four resistors in parallel, find the value of the four of them in parallel with R_5. When working out a problem of this type, sometimes it is worthwhile to look for groups that will work out in even numbers before starting the problem, and this will save some work. For example, look at the five resistors in parallel in Fig. 14.

In this example, if we group together R_1 and R_2 we get

$$R_t = \frac{12 \times 24}{12 + 24}$$

$$R_t = \frac{288}{36} = 8 \text{ ohms}$$

Therefore we can consider R_1 and R_2 as a single 8-ohm resistor. Now notice that the value of R_4 is 8 ohms, so let us consider the parallel combination of R_1 and R_2 which is equal to 8 ohms, with R_4. We can use the formula to find the value, but if we remember that the value of equal resistors in parallel is equal to one half the resistance of either resistor

ALSO! → $\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots \frac{1}{R_N}$

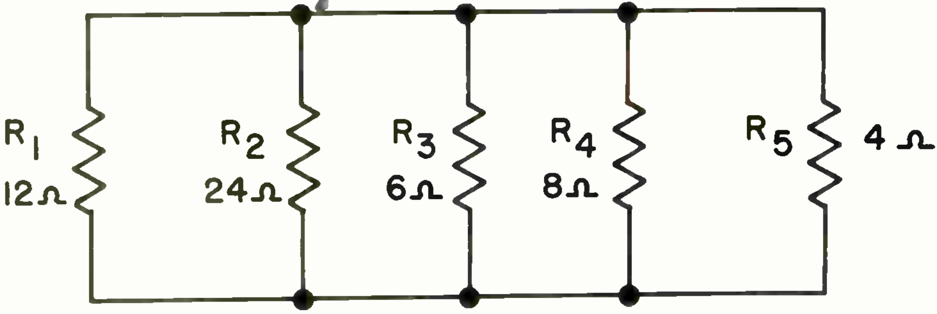


Fig. 14. Five resistors in parallel.

we know immediately that the parallel combination is equal to 4 ohms.

Now we consider the parallel combination of R_1, R_2 and R_4 as a single 4-ohm resistor. We now find the value of it in parallel with R_5, which is also a 4-ohm resistor, and we know immediately that the value of this combination is 2 ohms. Now we know that R_1, R_2, R_4 and R_5 in parallel have a resistance of 2 ohms. Now all we need to do is find the value of 2 ohms in parallel with R_3, which is 6 ohms, and we use the formula to do this.

$$R_t = \frac{6 \times 2}{6 + 2}$$

$$R_t = \frac{12}{8} = 1.5 \text{ ohms}$$

You can work this problem out by grouping other resistors together first; try working it out grouping R_1 and R_3 together first, get the value of this combination, and next get the value of R_2 and R_4 in parallel. After you have done this decide for yourself which two you should group together next to make your calculation as easy as possible, and then find the value of this combination in parallel with the remaining resistance to see if you get the same answer of 1.5 ohms.

Before leaving this section on resistors in parallel, there are a few other important points that you should notice.

Notice that in each of the examples given in Figs. 11, 12, 13 and 14 the total resistance of the resistors in parallel is less than the resistance of the smallest resistor. This is always true in a parallel circuit. Whenever you connect two or more resistors in parallel, the resistance of the parallel combination will always be less than the resistance of the smallest resistor, regardless of the value of the resistors connected in parallel.

If you look back at Fig. 11 you can see why this is so. You know from Ohm's Law that

$$R = \frac{E}{I}$$

Whenever you connect a resistor across a battery, a current flows in the circuit and for a given value of voltage a certain current will flow. If the voltage remains constant in the circuit and the current increases, then the resistance must get smaller. Whenever you connect a second resistor across a first resistor the current increases because there is a second path through which current can flow. Therefore when you connect two resistors in parallel and the value of E remains the same, the value of I increases and therefore the value of R must get smaller.

Sometimes when two resistors are connected in parallel, one resistor is so much larger than the other that practically all of the current flowing in the circuit flows through the smaller resistor. For example, if a 1000-ohm resistor is connected in parallel with a 1-ohm resistor and the two are connected across a battery, one thousand times as much current will flow through the 1-ohm resistor as through the 1000-ohm resistor. The current flowing through the 1000-ohm resistor would be such a small percentage of the total current flow in the circuit that it could be ignored. For all practical purposes, the total resistance of the two resistors in parallel can be considered as 1 ohm.

In most electronic circuits exact calculations are not necessary because most of the parts have a reasonable tolerance. If we want to get an approximate value of a group of resistors connected in parallel, we can usually ignore any resistor that is more than ten times larger than the smallest resistor in the circuit. Thus, if you had a 2-ohm, a 4-ohm and a 50-ohm resistor connected in parallel and wanted to get the approximate resistance of the parallel combination, you could ignore the 50-ohm resistor because it is more than ten times the resistance of the 2-ohm resistor. The value of the resistance you would obtain by ignoring this resistor would not differ appreciably from the value you would obtain if you did not ignore it and found the exact resistance of the parallel combination. Keep this in mind, because sometimes it will save you some unnecessary work if you are interested in finding only the approximate resistance of a group of resistors in parallel and do not need to know the exact resistance.

VOLTAGE AND CURRENT IN PARALLEL CIRCUITS

In a series circuit, the current is the same in all parts of the circuit. The voltage drop across a part in a series circuit depends upon the resistance of the part.

In a parallel circuit we have essentially the opposite situation. In a parallel circuit, since the parts are connected directly across each other, the voltage across each part must be the same. This means that the voltage across all the parts in a parallel circuit is the same. You cannot have two parts connected in parallel and have unequal voltages across them.

The current that will flow through each part in a parallel circuit will depend upon the voltage applied across the parallel circuit and the resistance of the part. Since each branch of a parallel circuit can have a different resistance, then we can have a different current flowing in each branch of the circuit.

Therefore, you should remember that in a parallel circuit the voltage will be the same across each part in the circuit. The current through each branch of the circuit will depend on the resistance of the branch. The highest current will flow through the branch having the lowest resistance and the lowest current will flow through the branch having the highest resistance. The total current flowing in a parallel circuit is equal to the sum of the currents flowing through the individual branches.

SUMMARY

Parallel circuits are widely used in electronic equipment and you will see many examples of them later in your course.

There are several important things you should remember about resistors in parallel. Remember that when two or more resistors are connected in parallel the total resistance of the parallel combination will always be less than the resistance of the smallest resistor.

When two equal value resistors are connected in parallel, the total resistance will be equal to one half the resistance of either resistor. If three equal value resistors are connected in parallel, the total resistance will be one third the resistance of either resistor, and if four equal value resistors are connected in parallel, the total resistance will be one quarter the resistance of any one resistor.

We can find the resistance of two resistors connected in parallel by using the formula

$$R_t = \frac{R_1 \times R_2}{R_1 + R_2}$$

In a parallel circuit the voltage across all branches of the parallel circuit will be equal. The current flow in each branch of the circuit will depend upon the resistance of the circuit, the highest current will flow through the lowest resistance branch, and the lowest current will flow through the highest resistance branch.

Now answer the self-test questions on this section of the lesson; they will be a good review for you.

SELF-TEST QUESTIONS

(l) What is the total resistance of two 50-ohm resistors connected in parallel?

(m) What is the total resistance of a 24-ohm resistor connected in parallel with a 12-ohm resistor?

(n) If a 3-ohm and a 4-ohm resistor are connected in parallel across a 12-volt battery, what will be the total current flow from the battery, and what will be the total resistance of the 3-ohm and 4-ohm resistance in parallel?

(o) If two 8-ohm resistors are connected in parallel with a 1000-ohm resistor, what will be the resistance of the parallel combination?

(p) If two resistors, R_1 and R_2, are connected in parallel across a battery and a current of 1 amp flows through R_1 and a current of 2 amps flows through R_2, which is the larger resistor? How much larger is it than the other resistor?

(q) If a 5-ohm resistor and a 10-ohm resistor are connected in parallel across a battery, and the current through the 5-ohm resistor is 2 amps, find the voltage across the 10-ohm resistor and the current that will flow through it. What is the total current flowing in the circuit? What is the battery voltage?

(r) Complete the following statement: in a parallel circuit, the voltages across all branches of the parallel circuit will be ________.

(s) In a parallel circuit, the current through each branch of a parallel circuit will depend upon ________ of the branch. The largest current will flow through the branch having the ______ resistance, and the smallest current will flow through the branch having the ________ resistance.

Series-Parallel Circuits

In radio and television transmitters and receivers as well as in all types of electronic control devices you will find many series circuits and many parallel circuits. However, in addition to these two types of circuits you will find many circuits that are combinations of series and parallel circuits. These types of circuits are called series-parallel circuits.

An example of a series-parallel circuit is shown in Fig. 15. In this circuit, the two resistors, R_2 and R_3, are connected in parallel. Any current flowing in the circuit, when it reaches the junction of R_2 and R_3 has two paths through which it can flow. Part of the current can flow through R_2 and part of it can flow through R_3. The exact percentage of the current that will flow through each resistor will depend upon their relative size. If the resistors are equal, half the current will flow through R_2 and half of it will flow through R_3. On the other hand, if one of the resistors is much larger than the other, most of the current will flow through the smaller resistor and only a small percentage of the current will flow through the larger resistor.

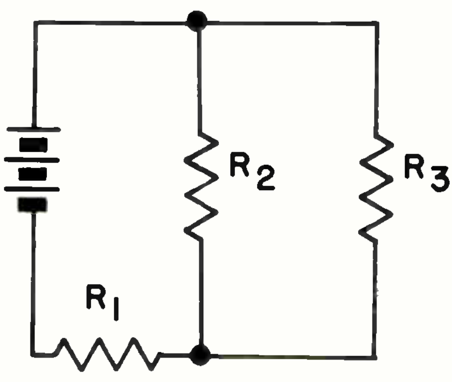


Fig. 15. A series-parallel circuit.

The two parallel resistors R_2 and R_3 are connected in series with the resistor R_1 and with the battery. Thus, part of the circuit is a simple series circuit and part of it is a parallel circuit.

In a circuit of this type, electrons leave the negative terminal of the battery and flow through the resistor R_1. The current divides after it flows through R_1, and part flows through R_2 and part flows through R_3. Then the current joins at the junction of the two resistors and flows back to the positive terminal of the battery and through the battery back to the negative terminal.

In solving for values of current, voltage and resistance in a series-parallel circuit, we use the rules

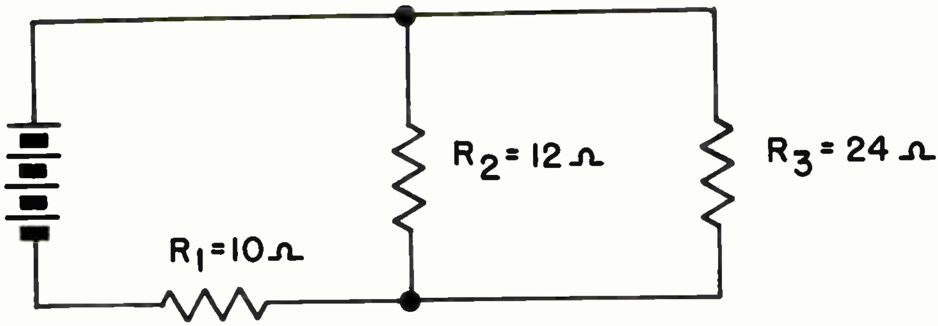


Fig. 16. A series-parallel circuit.

that apply to a series circuit for the series part of the circuit, and the rules that apply to a parallel circuit for the parallel part of the circuit. Actually, it is no more difficult to work with this type of circuit and find voltages, current etc., than it is with a simple series or a simple parallel circuit. However, there is a certain order in which you must proceed when working with circuits of this type. Once you become familiar with this type of circuit, you will see that it is not as complicated as it might look, and also learning about series-parallel circuits will make simple series and simple parallel circuits that much easier for you.

Now let us go ahead and study this type of circuit in detail.

RESISTANCE IN SERIES-PARALLEL CIRCUITS

In Fig. 16 we have shown a series-parallel circuit similar to the one shown in Fig. 15. In this example, R_1, which is called a series resistor, is equal to 10 ohms. R_2 and R_3 are called the parallel resistors. R_2 has a resistance of 12 ohms and R_3 has a resistance of 24 ohms. Now, we want to find the total resistance in the circuit. In other words, we want to find out how much resistance there is in the circuit limiting the current flow from the battery.

To find the total resistance of a series-parallel circuit of this type, the first step is to find the resistance of the parallel branch. To do this, we use the formula

$$R_p = \frac{R_2 \times R_3}{R_2 + R_3}$$

and we substitute the values of 12 ohms for R_2 and 24 ohms for R_3 and get

$$R_p = \frac{12 \times 24}{12 + 24}$$

$$R_p = \frac{288}{36} = 8 \text{ ohms}$$

Now, we know that the total resistance of R_2 in parallel with R_3 is 8 ohms. Insofar as the circuit is concerned, these two resistors could be replaced by a single 8-ohm resistor, as shown in Fig. 17. Notice the symbol we have used to indicate R_2 in parallel with R_3. The two parallel lines between R_2 and R_3 mean "in parallel with".

Once we replace the parallel combination of R_2 and R_3 with the equivalent resistance, we have a simple series circuit. As you will remember, the total resistance of a series circuit is equal to the sum of the individual resistances. In this case, the total resistance will be 10 + 8 = 18 ohms. This is the total resistance of the equivalent circuit shown in Fig. 17, and it is also the total resistance seen by the battery in the series-parallel circuit shown in Fig. 16.

So far the circuits we have been dealing with have been very simple series-parallel circuits. In some circuits we will have several series resistors and several parallel branches. The parallel branches may consist of any number of resistors. A more complicated series-parallel circuit is shown in Fig. 18. To find the total resistance seen by the battery in this circuit, we will use exactly the same method as we used to find the total resistance in the circuit shown in Fig. 16. The first thing we will do is find the effective resistance of the parallel branches and then redraw the circuit replacing the parallel branches

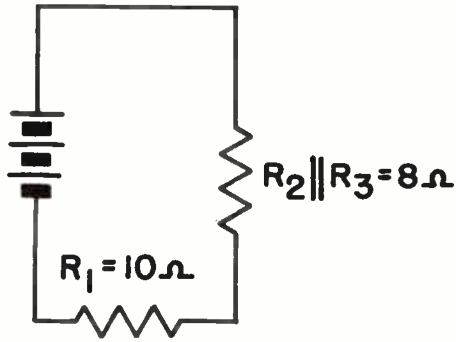


Fig. 17. Series equivalent of circuit of Fig. 15.

by single series resistors and then find the total series resistance. Now let us go through this problem step by step.

Notice that in the circuit shown in Fig. 18 we have two series resistors, R_1 and R_5. We also have two parallel branches. One parallel branch is made up of the three resistors R_2, R_3 and R_4, and the other

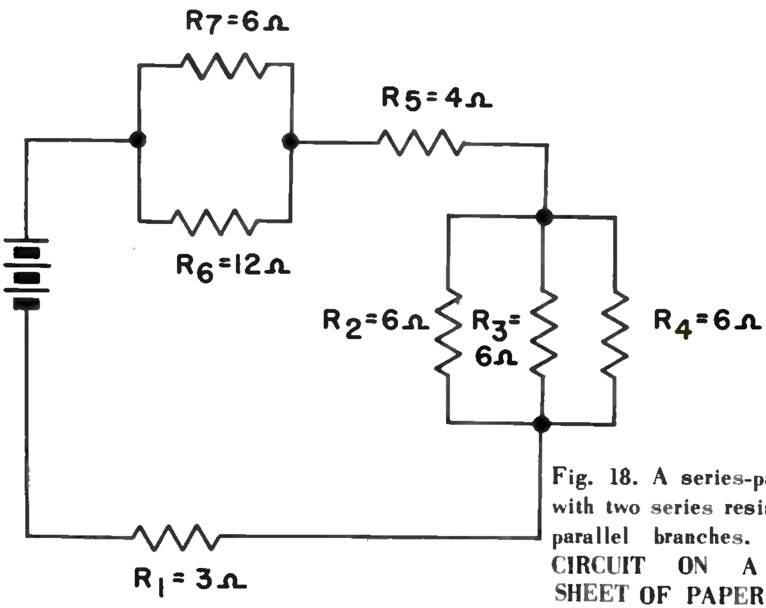


Fig. 18. A series-parallel circuit with two series resistors and two parallel branches. COPY THIS CIRCUIT ON A SEPARATE SHEET OF PAPER FOR QUICK REFERENCE.

parallel branch is made up of the two resistors R_6 and R_7.

Our first step in solving this problem is to find the resistance of the parallel branches. Let us take the branch made up of the three resistors first. If you will remember from the preceding section, we mentioned that when two equal value resistors are connected in parallel, the total resistance is equal to one half the resistance of either resistor. When three equal value resistors are connected in parallel the total resistance is one third the value of one of the resistors. From this we know immediately that the total resistance of the three six-ohm resistors in parallel is 2 ohms.

However, you might not remember this rule, or the resistors might not be of equal values. In this case you simply use the formula and find the resistance of two of the resistors in parallel. This would give you

$$R = \frac{6 \times 6}{6 + 6}$$

$$R = \frac{36}{12} = 3 \text{ ohms}$$

Now this gives you the value of two of the resistors in parallel and you combine this value with the remaining resistor to get the total resistance of the combination. Thus

$$R_t = \frac{3 \times 6}{3 + 6}$$

$$R_t = \frac{18}{9} = 2 \text{ ohms}$$

Therefore the resistance of the parallel branch made up of the three 6-ohm resistors is 2 ohms. As far as the circuit is concerned, we can replace these three resistors by a single 2-ohm resistor.

Now let us go ahead and find the value of the 12-ohm and 6-ohm resistors in parallel. Again, using our formula we get

$$R_t = \frac{12 \times 6}{12 + 6}$$

$$R_t = \frac{72}{18} = 4 \text{ ohms}$$

Therefore R_6 and R_7 can be replaced by a single 4-ohm resistor without affecting the total resistance of the circuit.

Now we can redraw the circuit shown in Fig. 18 as we have redrawn it in Fig. 19. In place of the three parallel resistors we have inserted a single 2-ohm resistor, and in place of resistors R_6 and R_7 we have inserted a 4-ohm resistor. Now we have a simple series circuit, and to get the total resistance of the circuit we simply add the resistance of the individual resistors. Therefore the total resistance in the circuit is

$$R_t = 3 + 2 + 4 + 4 = 13 \text{ ohms}$$

This same procedure can be used for solving any series-parallel circuit configuration. In fact, the circuit shown in Fig. 18 could consist of four parallel branches in series. There could be a resistor in parallel with R_1 and another resistor in parallel with R_5. In this case you would have a series circuit made

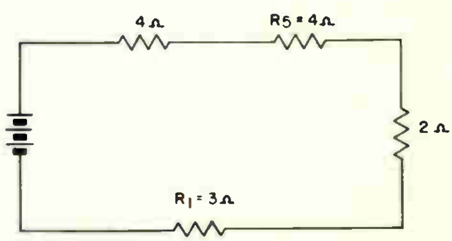


Fig. 19. Series equivalent of circuit of Fig. 18.

up of four parallel branches connected in series. To solve a circuit of this type you use the same procedure. Find the total resistance of each parallel branch, and then draw the equivalent circuit as we did in Fig. 19 and then total the resistance of the parallel branches to get the total resistance in the circuit.

Now let us see how the current behaves in a series-parallel circuit.

CURRENT IN SERIES-PARALLEL CIRCUITS

In a series-parallel circuit such as shown in Fig. 18, the current must be the same in every series part of the circuit. By this we mean that a certain current leaves the negative terminal of the battery and flows through R_1. Then the current reaches the parallel combination of R_2, R_3 and R_4. The total current flowing through this parallel combination must be equal to the current leaving the negative terminal of the battery and the current flowing through R_1. The same current must flow through R_5 and the same current must flow through the parallel combination of R_6 and R_7.

To help us see better what is happening to the current in the circuit, let's assume that the battery voltage in Fig. 18 is 78 volts. We picked this value because it will enable us to avoid any decimal divisions.

We have already found that the total resistance of the circuit - that is, the resistance that the battery sees - is 13 ohms. Therefore, we can use this to find the current that will flow in the circuit. Using the formula

$$I = \frac{E}{R}$$

and substituting 78 volts for E and 13 ohms for R we have

$$I = \frac{78}{13} = 6 \text{ amps}$$

This means that the current leaving the negative terminal of the battery is 6 amps. Since R_1 is in series with the battery, the current flowing through R_1 must also be 6 amps. The current flowing through the parallel combination of R_2, R_3 and R_4 must also be 6 amps. However, there are three paths here through which the current can flow. Since the three resistors are of equal value, then one third of the current will flow through each branch. This means that 2 amps will flow through R_2, 2 amps through R_3 and 2 amps through R_4.

The current reaching R_5 must flow through this resistor and since the series current is 6 amps, the current flow through R_5 must be 6 amps.

R_6 and R_7 form a parallel branch. Part of the current will flow through R_6 and part of it will flow through R_7. Since R_6 is twice the value of R_7, then the current flow through R_6 will be only half the current flowing through R_7. This means that one third of the current will flow through R_6 and two thirds of it will flow through R_7. Therefore the current flow through R_6 will be 2 amps, and the current flow through R_7 will be 4 amps. The current flowing into the positive terminal of the battery will be 6 amps, and the current flow through the battery itself from the positive terminal to the negative terminal will also be 6 amps.

Now notice how this circuit has

obeyed the laws set down both for series circuits and for parallel circuits. We stated that in a series circuit, the current is the same in all parts of the circuit. R_1 is a part of the series circuit and the current through it is 6 amps. R_2, R_3 and R_4 in parallel form a part of the series circuit and the current through the three in parallel is 6 amps. The current divides because there are three paths, but the total current flowing through the three paths is 6 amps. R_5 does not have any other resistors in parallel with it and therefore the entire current of 6 amps flows through it. R_6 and R_7 in parallel form a part of the series circuit and the total current flowing through the two of them is 6 amps; part of it flows through R_6 and part of it through R_7. The entire 6 amps also flows through the battery. Thus we can say that the current flow in each part of this series circuit is 6 amps.

The current flow also obeys the law of parallel circuits. Whenever the current reaches a parallel branch it divides, and part of it flows through each branch of the circuit. Since R_2, R_3 and R_4 are of equal value, the same current flows through each of the three resistors. In this case, 2 amps flow through each resistor so that the total current flowing in this part of the series circuit is 6 amps. In the case of R_6 and R_7 in parallel, the current again divides and flows through the two resistors in proportion to their resistance. R_6 is twice the resistance of R_7 and therefore only half as much current flows through R_6 as R_7. 2 amps flow through R_6 and 4 amps through R_7. This gives us a total current of 6 amps, once again through this part of the series circuit.

Now that we have looked into resistance and current in series-parallel circuits, let us see what happens to the voltage in these circuits.

VOLTAGE IN SERIES-PARALLEL CIRCUITS

You will remember that in a series circuit, the sum of the voltage drops across the branches of the series circuit is equal to the applied voltage. You will also remember, that in a parallel circuit you have the same voltage across all the parts in parallel. Now let us see what happens in a series-parallel circuit. Let's use the circuit in Fig. 18 so we don't have to figure out the total resistance again, and once again let's assume that the battery voltage is 78 volts. We know that this will cause a current of 6 amps to flow in the circuit.

With a current of 6 amps flowing through R_1, we can find the voltage drop across it. We will refer to this voltage as E_1, the voltage across R_2 as E_2 and so on.

$$E_1 = 6 \times 3 = 18 \text{ volts}$$

We know that the current flowing through R_2 is 2 amps, and using this we can find the voltage across R_2.

$$E_2 = 2 \times 6 = 12 \text{ volts}$$

Since $R_2 = R_3 = R_4$, and the current flowing through each of these resistors is 2 amps, they must have the same voltage drop across each of the resistors. This bears out what we have said before about parts connected in parallel – the voltage drop across them must be the same.

If we stop here and look at Fig. 19, we see that we found the equivalent resistance of R_2, R_3 and R_4 connected in parallel was 2 ohms. In this circuit, the current flowing through this equivalent resistance is 6 amps, and the voltage across it will be 12 volts. Therefore we see that no matter which way we work, either using the total current across the equivalent resistance or the actual current through one branch of the parallel circuit, we should get the same value of voltage across the parallel circuit.

The voltage across R_5 will be

$$E_5 = 6 \times 4 = 24 \text{ volts}$$

The voltage across R_6 will be

$$E_6 = 2 \times 12 = 24 \text{ volts}$$

We know that the current through R_7 is 4 amps and so we can also calculate the voltage across this resistor to see that we have the same voltage as across R_6. The voltage across R_7 is

$$E_7 = 4 \times 6 = 24 \text{ volts}$$

Now, let us add the voltage drops around the circuit to see what the total voltage drop is. The voltage drop across R_1 is 18 volts. The voltage drop across the parallel combination of R_2, R_3 and R_4 is 12 volts. The voltage drop across R_5 is 24 volts and the voltage drop across the parallel combination of R_6 and R_7 is also 24 volts. Now adding these voltages we will get the total voltage drop in the circuit

$$E_t = 18 + 12 + 24 + 24 = 78 \text{ volts}$$

The series-parallel circuit thus obeys the law we stated for the voltage drops in a series circuit. The sum of the voltage drops across each of the series branches in a series-parallel circuit is equal to the source voltage. The voltage drop across each of the parts in a parallel branch is equal. Thus we can see that the rules we have set up for voltages in series and in parallel circuits apply also to series-parallel circuits.

Now, if you plan to become a radio-TV service technician, this is as far as you have to go with series-parallel circuits. However, if you want to go ahead and get your FCC license, or if you plan on working in industry, you should be able to solve some simple problems in series-parallel circuits. You will have to solve these problems to get your FCC license. Often when applying for positions in industry, examinations are given and frequently these examinations have one or two problems involving solutions of parallel series and series-parallel circuits. For the technician who wants to get his FCC license or wants to work in industry, and for the radio-TV service technician who wants to learn more about series-parallel circuits we will go through a typical problem in the next section. Remember, if your interest is radio-TV servicing only, you do not have to go through this section unless you want to.

PROBLEMS IN SERIES-PARALLEL CIRCUITS

In Fig. 20 we have shown a series-parallel circuit that is only slightly more complicated than the one shown in Fig. 18. In this circuit you are given the values of all the resistors

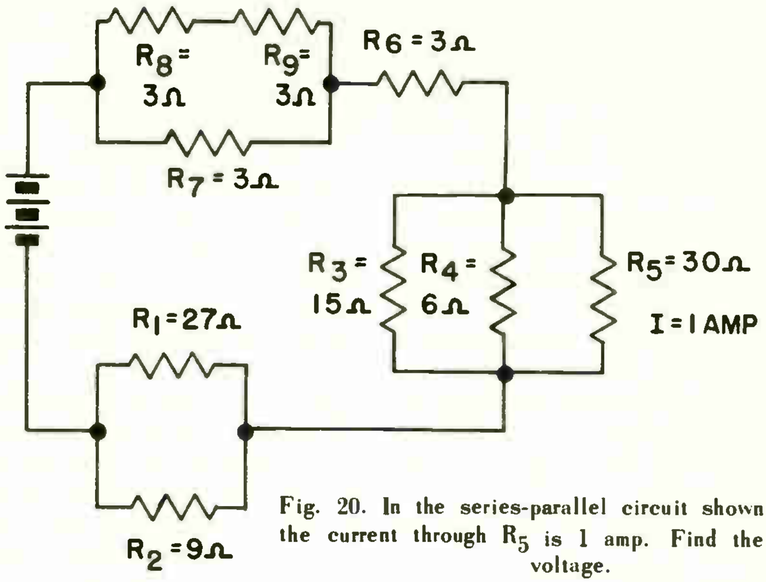


Fig. 20. In the series-parallel circuit shown above, the current through R_5 is 1 amp. Find the source voltage.

and also the information that the current flowing through R_5 is 1 amp. The problem is to find the source voltage.

At first glance, you might think that this is impossible, but actually it is not nearly as difficult as it might appear at first. Let us look at the information we have about R_5.

We know that the resistance of R_5 is 30 ohms, and that the current flowing through it is 1 amp. From this we can calculate the voltage across R_5 using Ohm's Law. We use the formula

$$E = IR$$

and for I substitute 1 amp and for R, 30 ohms, and we will get

$$E = 1 \times 30 = 30 \text{ volts}$$

Since R_3 and R_4 are in parallel with R_5, this means that the voltage across R_3 and the voltage across R_4 is also 30 volts. Now that we know the voltage across these two resistors, and their value, we can find the current flowing through them. We will use I_3 for the current through R_3 and I_4 for the current through R_4 and so on throughout the problem.

$$I_3 = \frac{30}{15} = 2 \text{ amps}$$

$$I_4 = \frac{30}{6} = 5 \text{ amps}$$

Now we know that the current flowing through R_3 is 2 amps, through R_4 is 5 amps, and the current flowing through R_5 is 1 amp. Thus the total current flowing through the three parallel branches is

$$I_t = 2 + 5 + 1 = 8 \text{ amps}$$

Since this combination of three resistors forms one branch in series circuit, now we know that the series current must be 8 amps. We can immediately calculate the voltage drop across R_6 using this information.

$$E_6 = 8 \times 3 = 24 \text{ volts}$$

Now we know that the voltage drop across R_6 is 24 volts and the voltage drop across the parallel combination of the three resistors is 30 volts. We know that the source voltage will be equal to the sum of the voltage drops around the series circuit so if we can find the voltage drop across the parallel combination of R_1 and R_2 and also the voltage drop across the parallel combination of R_7 in parallel with R_8 and R_9, we'll be able to find the sum of the voltage drops and hence the source voltage.

We can find the voltage drop across the parallel combination of R_1 and R_2 quite easily. Notice that R_1 is three times as large as R_2. This means that the current flowing through R_2 will be three times the current flow through R_1. In other words, if the current flow through R_1 is 1 amp, then the current flow through R_2 would be 3 amps which would give us a total of 4 amps flowing through the two resistors. However, we know that the total current flow in the circuit is 8 amps, therefore the current flowing through R_2 must be 6 amps and the current flowing through R_1 must be 2 amps.

Sometimes you can't divide the current flow this easily, but we can still find what the voltage drop across the two resistors must be by finding the resistance of the parallel combination. To do this, we substitute the values of R_1 and R_2 in our parallel formula

$$R = \frac{27 \times 9}{27 + 9}$$

$$R = \frac{243}{36} = 6.75 \text{ ohms}$$

Now we know that the resistance of R_1 and R_2 in parallel is 6.75 ohms and that the current flowing through the parallel combination is 8 amps; we can find the voltage drop across this combination.

$$E = 8 \times 6.75 = 54 \text{ volts}$$

If you noticed that the current flow through R_2 must be 6 amps and the current flow through R_1 must be 2 amps, then you can use either current flow along with the appropriate resistor to get the voltage drop across the parallel combination.

$$E_1 = 2 \times 27 = 54 \text{ volts}$$

and if you use R_2 in the current flow through it, you would get

$$E_2 = 6 \times 9 = 54 \text{ volts}$$

Notice that no matter how you work it out, either finding the total equivalent resistance and using the total current, or using the resistance of either branch and the current flow through it, you always get the same voltage drop across the parallel combination or R_1 and R_2.

Now all we need to do is find the equivalent resistance of R_7 in parallel with R_8 and R_9, and we can find the voltage drop across this combination. The first step is to find the resistance of R_8 in series with R_9. This is simply 3 + 3 = 6 ohms. Now we have a total of 6 ohms in parallel with 3 ohms; we want to find the resistance of this parallel combination. Again, we use our formula for parallel resistors and get

$$R = \frac{3 \times 6}{3 + 6}$$

$$R = \frac{18}{9} = 2 \text{ ohms}$$

Since the parallel combination of R_7 in parallel with R_8 and R_9 has a resistance of 2 ohms, and the current flowing through this combination is 8 amps, we can find the voltage drop across the parallel combination

$$E = 8 \times 2 = 16 \text{ volts}$$

Now we have the voltage drop across each branch in the series-parallel circuit and all we need to do is add these voltage drops - they must be equal to the source voltage. Therefore the total voltage is equal to

$$E_t = 54 + 30 + 24 + 16 = 124 \text{ volts}$$

This is typical of the problems involving the series-parallel circuits that you might have to work out either on an FCC examination or on an examination for employment in industry. Notice that the problem is not difficult; you just start using Ohm's Law in a part of the circuit where you have two of the three quantities, E, I or R, and then find the third. You use this information to learn more about the circuit and work one step at a time. In the circuit shown in Fig. 20 we were able to determine the current flowing in each part of the circuit, the voltage drop in each branch of the circuit, and the total voltage applied to the circuit. If we wanted to, we could also add the resistance of the individual branches of the circuit and find the total resistance of the circuit; or, since we know the total voltage in the circuit we could simply divide this by the current to get the total resistance in the circuit. Solving problems of this type are really not any more difficult than solving simple Ohm's Law problems; they are simply longer because there are more steps involved.

SUMMARY

Series-parallel circuits are important because you will run into them in all types of electronic equipment. In a series-parallel circuit you will have a voltage source in series with a combination of series and parallel components. You may have one series branch and one parallel branch or you may have several of each. In the parallel branches there may be any number of components.

In a series-parallel circuit the total current flow is the same in all series branches of the circuit. In the parallel branches, the sum of the currents in the individual branches must be equal to the series current flow. The voltage drop across all the components in the parallel branch is the same and the voltage drop across the series components will depend upon the resistance of the component and the total current flowing in the circuit.

To find the total resistance in a series-parallel circuit, you reduce the parallel branches to the equivalent series resistance, and then add the series branches with the equivalent resistance of the parallel branches. The total current flow in the circuit can be determined from the source voltage divided by the total resistance of the circuit and the voltage drop across individual parts in the circuit can be found by using Ohm's Law.

Solving problems involving series-parallel circuits is simply a matter of taking one step at a time, making simple applications of Ohm's Law until the entire circuit is solved.

The following self-test questions are designed to help you with series-parallel circuits. If you find that you are having difficulty with one of the problems, be sure to go back and review. Try to work the problem through on your own, but if you find you can't, look for help at the back of the book - the solutions to the problems are given there. If you have to go to the back of the book to try to find out how to do a problem, make sure you understand that problem before tackling the next one; the chances are that if you will do this, then you will be able to work the next problem by yourself.

SELF-TEST QUESTIONS

(t) Draw a series-parallel circuit containing a battery and three resistors in which R_1 and R_2 are in parallel and connected to the negative terminal of the battery, and R_3 is a series resistor and connected to the positive terminal of the battery.

(u) If R_1 = 20 ohms and R_2 = 30 ohms, and R_3 = 12 ohms, find the total resistance in the circuit.

(v) Using the values of R_1, R_2 and R_3 for the preceding problem, if the battery voltage is 48 volts, find the voltage drop across each resistor in the circuit, and the current flow through each resistor.

(w) If in a circuit like the one shown in Fig. 15, R_1 = 5 ohms, R_2 = 10 ohms and R_3 = 10 ohms, what is the source voltage if the current through R_2 is 1 amp.

(x) Find the total resistance in a series-parallel circuit like the one shown in Fig. 18 when R_1 = 5 ohms, R_2, R_3 and R_4 are each equal to 12 ohms, R_5 = 6 ohms, R_6 = 48 ohms and R_7 = 24 ohms.

(y) Complete the following statement: In a series-parallel circuit, the current is the ________ in each series branch of the circuit.

(z) Complete the following statement: In a parallel branch of a series-parallel circuit, the sum of the currents through the branches of the parallel circuit is ________ to the total current flowing in the series part of the circuit.

ANSWERS TO SELF-TEST QUESTIONS

(a) A series circuit is a circuit in which the voltage source and the parts are connected, so that current leaving the negative terminal of the voltage source flows through first one part and then another to the positive terminal of the voltage source and then through the source back to the negative terminal.

(b) To measure the voltage of a battery with a dc voltmeter, you connect the negative terminal of the voltmeter to the negative terminal of the battery, and the positive terminal of the voltmeter to the positive terminal of the battery.

(c) See Fig. 1 and Fig. 3. Either figure is correct; they are the same circuit except in Fig. 3 the resistor values are given.

(d) The polarity of the voltage drops across the resistors is indicated in Fig. 8.

(e) The total resistance in the circuit will be equal to the resistance of the sum of the resistors. Therefore

$$R_t = 3 + 4 + 5 = 12 \text{ ohms}$$

(f) To find the voltage across any one of the resistors you simply use Ohm's Law. You know that

$$E = IR$$

You know the value of each of the resistors and the current will be the same through all resistors so you can find the voltage drop across each one.

$$E_{r1} = 2 \times 3 = 6 \text{ volts}$$

$$E_{r2} = 2 \times 4 = 8 \text{ volts}$$

$$E_{r3} = 2 \times 5 = 10 \text{ volts}$$

(g) You can find the source voltage in two ways. You can add the individual voltage drops you obtained in the preceding section and you will find that the source voltage is

$$6 + 8 + 10 = 24 \text{ volts}$$

You can also get the source voltage from the current times the total resistance, which you found to be 12 ohms. Using this you get

$$E = 2 \times 12 = 24 \text{ volts}$$

(h) You know that the sum of the voltage drops in a series circuit is equal to the source voltage. In this example, the source voltage is 35 volts. The voltage drops across the three resistors are 5 volts, 7 volts and 10 volts. Adding these three voltage drops, we get 22 volts. Therefore the voltage drop across the remaining resistor must be 13 volts.

(i) This problem is exactly the same as the preceding problem. In a series circuit, the sum of the voltage drops is equal to the source voltage. The source voltage is 250 volts, and therefore the voltage drops in the series circuit must add up to 250 volts. The voltage drop across R_1 is 3 volts, and the voltage drop across R_2 is 125 volts. Therefore the sum of the voltage drops across R_1 and R_2 is 128 volts. The remainder of the 250 volts must be dropped between the plate and cathode of the vacuum tube. Therefore the voltage drop across the tube must be 122 volts.

(j) This problem is the same as the preceding problem. The voltage drop across R_1 plus the voltage drop across R_2 is equal to 3.8 volts. Since the battery voltage is 6 volts, the difference, which is 2.2 volts, must be the voltage dropped between the emitter and the collector of the transistor.

(k) This problem is a review of just about everything you have learned about series circuits. To find the total resistance in a series circuit you add the resistance of the individual resistors. Therefore

$$R_t = 120 + 150 + 180 + 120 + 330 = 900 \text{ ohms.}$$

To find the current flowing in the circuit you use Ohm's Law in the form

$$I = \frac{E}{R}$$

Substituting 18 volts for E and 900 ohms for R we get

$$I = \frac{18}{900}$$

This will involve a decimal division and so we multiply it by 1000 and get our current directly in milliamperes. Thus our problem becomes

$$I = \frac{18}{900} \times 1000$$

$$I = \frac{18000}{900}$$

Now we cancel two zeros above and below the line and then we have

$$I = \frac{180}{9} = 20 \text{ ma}$$

To find the voltage drop across each of the resistors we use the formula

$$E = IR$$

We will refer to the voltage drop across R_1 as E_1 and the voltage drop across R_2 as E_2 and so on.

$$E_1 = \frac{20}{1000} \times 120$$

Notice that we wrote the current as

$$\frac{20}{1000}$$

The current must be in amperes and we convert 20 milliamperes to amperes by dividing it by 1000. Now continuing with the problem,

$$E_1 = \frac{20 \times 120}{1000}$$

$$E_1 = \frac{2400}{1000}$$

$$E_1 = \frac{24}{10} = 2.4 \text{ volts}$$

$$E_2 = \frac{20}{1000} \times 150$$

$$E_2 = \frac{3000}{1000} = 3 \text{ volts}$$

$$E_3 = \frac{20}{1000} \times 180$$

$$E_3 = \frac{3600}{1000}$$

$$E_3 = \frac{36}{10} = 3.6 \text{ volts}$$

$$E_4 = E_1 = 2.4 \text{ volts}$$

$$E_5 = \frac{20}{1000} \times 330$$

$$E_5 = \frac{6600}{1000}$$

$$E_5 = \frac{66}{10} = 6.6 \text{ volts}$$

Notice that we did not calculate the voltage across R_4. R_4 is equal to R_1 and since in a series circuit the same current flows through the entire circuit then the voltage drop across R_4 must be equal to the voltage drop across R_1. Therefore calculating the voltage a second time would simply be a waste of time. As a matter of fact, sometimes in a problem of this type you can save yourself some work by looking up the resistor values. For, example, if one resistor happened to be twice that of the other, you would know immediately that the voltage drop across it was twice as high as the voltage drop across the first resistor. Similarly if one resistor was three times another you could expect three times as high a voltage drop across it.

(l) 25 ohms. The total resistance of two equal resistors connected in parallel is always one half the resistance of either resistor.

(m) 8 ohms. You find the resistance of the two resistors in parallel by using the formula

$$R_t = \frac{R_1 \times R_2}{R_1 + R_2}$$

and substituting 24 ohms and 12 ohms for R_1 and R_2 you get

$$R_t = \frac{24 \times 12}{24 + 12}$$

$$= \frac{288}{36} = 8 \text{ ohms}$$

(n) The current flow through each resistor can be found from Ohm Law's

$$I = \frac{E}{R}$$

The current through the 3-ohm resistor will be

$$I = \frac{12}{3} = 4 \text{ amps}$$

and the current through the 4-ohm resistor will be

$$I = \frac{12}{4} = 3 \text{ amps}$$

The total current flow will be the sum of these two currents or 7 amps. You can use this value of current and the voltage of 12 volts to find the resistance of the two resistors in parallel,

$$R = \frac{E}{I}$$

$$R = \frac{12}{7} = 1.7 \text{ ohms}$$

(o) 4 ohms. The resistance of two 8-ohm resistors in parallel will be 4 ohms, one half the resistance of either resistor. The 1000-ohm resistor is so large that we can simply ignore it because it will not affect the total resistance of the circuit appreciably. As a matter of fact, if you did consider it and calculated the value of the three in parallel you would find that the resistance worked out to be over 3.98 ohms. This is less than two parts in four hundred, an error which is so small it is insignificant, and it can be ignored.

(p) R_1 is the larger resistor because the smaller current flows through it. Since only half as much current flows through R_1 as through R_2, then R_1 must be twice the size of R_2.

(q) If the current through the 5-ohm resistor is 2 amps, then the voltage across it must be 10 volts. We use Ohm Law's in the form E = IR to determine this voltage. Since the two resistors are in parallel and connected across a battery we know that the battery voltage must be 10 volts, and the voltage across the 10-ohm resistor must be 10 volts. With 10 volts across the 10-ohm resistor, a current of 1 amp will flow through it. We would also know that a current of 1 amp must flow through the 10-ohm resistor since it is twice the size of the 5-ohm resistor, and therefore half the current would flow through it that flows through the 5-ohm resistor. The total current flowing in the circuit must be the sum of the current through the two resistors or 3 amps.

(r) In a parallel circuit, the voltage across all branches of the parallel circuit will be equal.

(s) In a parallel circuit, the current through each branch of a parallel circuit will depend upon the resistance of the branch. The largest current will flow through the branch having the lowest resistance, and the smallest current will flow through the branch having the highest resistance.

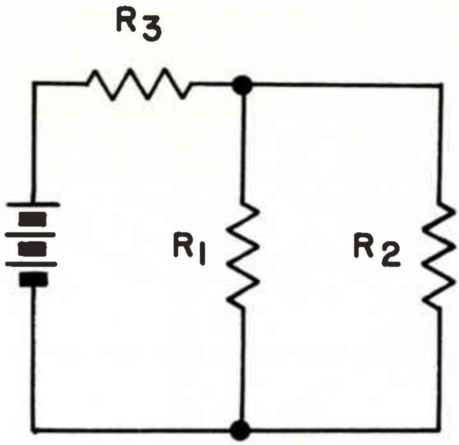


Fig. 21. Answer to Self-Test Question (t).

(t) See Fig. 21.

(u) 24 ohms. The total resistance of the parallel combination of R_1 and R_2 can be found from the parallel resistor formula. Substituting these values we get

$$R_t = \frac{20 \times 30}{20 + 30}$$

$$= \frac{600}{50} = 12 \text{ ohms}$$

This resistance is in series with R_3, which also has a resistance of 12 ohms, so the total resistance in the circuit is 24 ohms.

(v) With a voltage of 48 volts and a total resistance of 24 ohms, the total current flowing in the circuit will be

$$I = \frac{48}{24} = 2 \text{ amps}$$

This means that the voltage drop across R_3 will be

$$E = 2 \times 12 = 24 \text{ volts}$$

Therefore the voltage drop across the parallel combination of R_1 and R_2 must also be 24 volts. The current through R_1 must be

$$I = \frac{24}{20} = 1.2 \text{ amps}$$

The current through R_2 must be

$$I = \frac{24}{30} = .8 \text{ amps}$$

(w) 20 volts. If the current through R_2 is 1 amp, and the resistance of R_2 is 10 ohms, then the voltage across R_2 must be

$$E = 1 \times 10 = 10 \text{ volts}$$

Since R_2 and R_3 are in parallel, then the same voltage must be across R_3, and therefore the same current of 1 amp will flow through it, giving a total current of 2 amps through R_2 and R_3. This current of 2 amps must flow through R_1 and since R_1 has a resistance of 5 ohms, the voltage drop across it must be

$$E = 2 \times 5 = 10 \text{ volts}$$

Since the source voltage in any circuit is equal to the sum of the voltage drops around the circuit, the source voltage in this case must be 20 volts.

(x) The total resistance is 31 ohms. To solve this problem we must find the resistance of R_2, R_3 and R_4 in parallel and also the resistance or R_6 and R_7 in parallel. We can substitute an equivalent for these two parallel groups; then we have a simple series circuit and we can find the total resistance of this circuit simply by adding the individual resistances.

Since R_2, R_3 and R_4 are each 12-ohm resistors, then the total resistance of the three resistors in parallel will be one-third the resistance of any one of the resistances or one-third of 12 which is 4 ohms.

The resistance of the parallel combination of R_6 and R_7 can be found using the parallel resistance formula

$$R_t = \frac{48 \times 24}{48 + 24} = 16 \text{ ohms}$$

Now substituting 4 ohms for the parallel combination R_2, R_3 and R_4 and 16 ohms for the parallel combination of R_6 and R_7 we have a series resistance circuit. The total resistance of this circuit will be

$$R_t = 5 + 4 + 6 + 16 = 31 \text{ ohms}$$

(y) In a series-parallel circuit, the current is the same in each series branch of the circuit.

(z) In a parallel branch of a series-parallel circuit, the sum of the currents through the branches of the parallel circuit is equal to the total current flowing in the series part of the circuit.

Lesson Questions

Be sure to number your Answer Sheet B104-1.

Place your Student Number on every Answer Sheet.

1. What is the total resistance of three 15-ohm resistors connected in series?

2. What is the total resistance of three 15-ohm resistors connected in parallel?

3. A 3-ohm, a 5-ohm and a 4-ohm resistor are connected in series across a battery. The voltage drop across a 3-ohm resistor is 6 volts. What is the battery voltage?

4. If three resistors are connected in series across a 12-volt battery, and the voltage drop across one resistor is 3 volts and the voltage drop across the second resistor is 7 volts, what is the voltage drop across the third resistor?

5. If four resistors are connected in parallel and R_1 equals 12 ohms and R_2 equals 9 ohms, R_3 equals 24 ohms and R_4 equals 18 ohms, what is the resistance of the parallel combination?

6. R_1 and R_2 are two resistors of the same value connected in parallel. This parallel combination is connected in series with R_3, a 10-ohm resistor. The series-parallel network is connected to a 15-volt battery. If the voltage across R_3 is 10 volts, what are the values of R_1 and R_2?

7. If two resistors are connected in parallel, will the voltage drop across the two be (1) equal, (2) greater across the larger resistor, (3) greater across the smaller resistor?

8. Resistors R_1 and R_2 are connected in parallel. The resistance of R_1 is 4 ohms, and the current through it is 2 amps. The current through R_2 is .5 amps. Find the resistance of R_2.

9. Three 6-ohm resistors are connected in parallel. A fourth 6-ohm resistor is connected in series with the parallel combination. The series-parallel network is connected to a battery with the free end of the single 6-ohm resistor going to the positive terminal. Draw a schematic diagram of the circuit.

10. In the circuit of Question 9, if the battery voltage is 8 volts, what will the voltage across the three parallel-connected resistors be?

HOW TO CONCENTRATE

The secret of rapid progress with any course of study lies in being able to concentrate. If your mind wanders from study while you are alone in a quiet room, try moving to a noisy room, or try tuning in an all-musical program on the radio. Stay away from loud conversations, though.

If noise definitely bothers you, however, do your studying in a quiet location. If you feel too sleepy to concentrate, the room may be too hot. Open the windows, and put on a coat if necessary, for it is easier to study in a cool room. Sponge your face, neck and eyes with cold water.

If you have difficulty in understanding a subject, write down an explanation of it in your own words, or try outlining the subject and studying your outline. Underlining important words as you study is another aid to learning.

Try different studying techniques, until you find the one that gives you complete mastery of a subject in the shortest possible time. The faster you master each lesson, the sooner you'll complete the Course and be ready for a good job in electronics.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-204

ACHIEVEMENT THROUGH ELECTRONICS

HOW RESISTORS ARE USED

B105-1

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

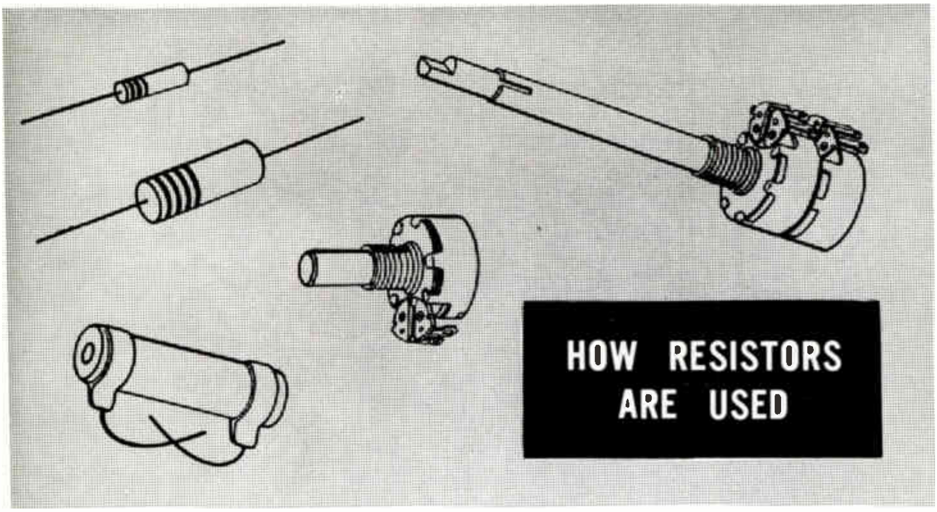
HOW RESISTORS ARE USED

B105-1

STUDY SCHEDULE

For each study step, read the assigned pages first at your usual speed, then reread slowly one or more times. Finish with one quick reading to fix the important facts firmly in your mind. Study each other step in the same way.

JDL3M1174 **1974 EDITION** **Litho in U.S.A.**



This is the first lesson in your NRI course that will be devoted almost entirely to one electronic part - resistors. In your next lesson you will study coils in detail, and in the following lesson you will study capacitors. These three parts - resistors, coils and capacitors - along with tubes and transistors are the most important parts in electronic equipment. You will find more resistors in electronic equipment than any other part. Therefore it is important that you learn how they are used so that you will be able to tell whether or not a particular circuit is working properly, and also so that you will be able to select a suitable replacement for a resistor in a circuit, when a replacement is needed.

Resistors, coils, and capacitors are important, not only because they are used so often in electronic equipment, but also understanding how these parts work will help you to understand how other parts work. For example, a transformer is nothing more than a group of coils wound on a single core. If you understand how a single coil works, then you will be much better prepared to understand how several of them work when used together to make up a transformer.

Resistors perform in essentially the same way in both ac and dc circuits. However, you will find when you study coils and capacitors that this is not true of these parts. Actually, coils and capacitors have little or no value in dc circuits. They are primarily used in ac circuits. Therefore, when you study coils and capacitors, you will be learning more about these parts, and we will also go into more detail about ac and how it acts in circuits where these parts are present.

WHY RESISTORS ARE IMPORTANT

Resistors are found in practically every piece of electronic equipment. As a technician you will have to replace many resistors. As we mentioned earlier, there are more resistors used in electronic equipment than any other parts.

Sometimes when you have to replace a resistor you will be able to refer to the schematic diagram of the equipment, find the value of the resistor and simply go ahead and install a new resistor in the circuit. However, you will find that on many occasions you will have to work on electronic equipment for which there is no diagram available and the original resistor may have been burned so badly that you will be unable to tell what its value was. Then, you will have to fall back on your knowledge of electronic circuits to decide what size resistor to use. What you will learn in this lesson will prepare you for this type of work.

There are many uses for resistors in electronic equipment. The various electrodes in a tube or transistor used in a piece of electronic equipment do not all require the same operating voltages. However, for economy, the required operating voltages must all be obtained from a single power supply. Resistors are used to drop the voltage to the correct value for the tube or transistor.

Resistors are used to isolate parts from each other, so that one will not interfere with the operation or the action of each other.

Special variable resistors called potentiometers are used to control the volume and tone in radio and TV receivers, to control the picture brightness and contrast in black and white and color TV receivers, and to adjust the tint and color saturation of the picture in color sets.

There is no end to the uses to which resistors are put in electronic equipment, and therefore it is extremely important that you understand how they are used.

TYPES OF RESISTORS

You already know that there are several types of resistors used in electronic equipment. The most frequently encountered resistor is the carbon resistor, which as we mentioned before is simply a carbon compound which is held together by a cement-type of binder. Carbon resistors are made chiefly in 1/2-watt, 1-watt and 2-watt sizes. Occasionally you will run into very small carbon resistors that are 1/3-watt resistors, but these are not used too often. You will learn more about the watt, which is the unit of power measurement, later in this lesson.

You will soon learn to recognize the wattage rating of a resistor by its size. The resistors shown in Fig. 1 are 1/2-watt, 1-watt and 2-watt carbon resistors. The body of the resistor in each case is drawn full size to give you an idea of how big each type of resistor is.

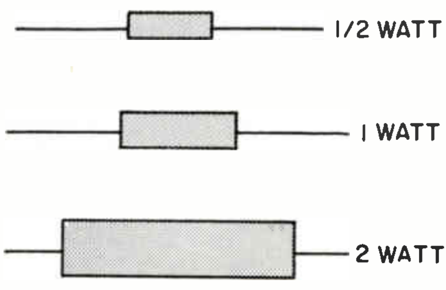


Fig. 1. Relative physical size of the three different wattage carbon resistors usually found in electronic equipment.

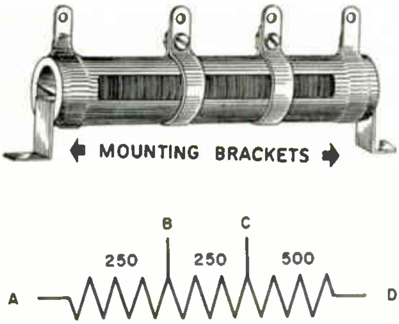


Fig. 2. A tapped wire-wound resistor and its schematic symbol.

In addition to carbon resistors you will run into wire-wound resistors. A wire-wound resistor is made by winding a resistance wire on a form. Wire-wound resistors are found with wattage ratings of 3 or 4 watts up to very high wattage ratings.

In some electronic equipment you will find tapped wire-wound resistors that look like the resistor shown in Fig. 2. The total resistance of this resistor is 1000 ohms. This is the resistance you would measure between terminals A and D. However, there are two taps on the resistor; these are the taps B and C. The resistance between terminal A and B is 250 ohms and the resistance between terminals B and C is 250 ohms. The remaining resistance between terminals C and D is 500 ohms. This type of resistor is usually quite a large resistor and in most cases it is mounted on the chassis by means of mounting brackets which hold it in place. The resistor is so large and heavy that if it were simply held in place by wires connected to the various terminals, the chances are that if the equipment received any jarring or bouncing the resistor would break loose from the wires.

Metal oxide resistors are also widely used in modern electronic equipment. These resistors can be made in higher wattage ratings than carbon resistors, and still have many of the advantages of the carbon resistors. The wire-wound resistors are made by winding wire in the form of a coil on an insulated support such as a ceramic type rod. Since the resistor is made of a coil of wire it takes on some of the characteristics of a coil. In some circuits this may be undesirable and in applications of this type a metal oxide resistor can be used. It is made by depositing a metal oxide film on a ceramic or glass tube or rod. The oxide film is in the form of a continuous path rather than a coil and therefore does not act like a coil.

You will also run into variable resistors such as shown in Fig. 3. The resistor shown in Fig. 3A is a wire-wound type of variable resistor and is usually called a rheostat. Notice that this type of resistor has only two terminals. Another rheostat is shown in Fig. 3B. This rheostat can either be a wire-wound control similar to the one shown in A or it might have a carbon element and a slider that rotates along the

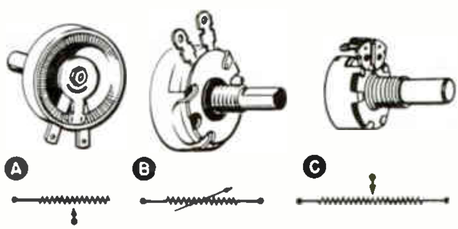


Fig. 3. Variable resistors and the symbols for them. A and B are rheostats, and C is a potentiometer.

carbon element to provide the required resistance between the two terminals. The variable resistor shown in Fig. 3C is called a potentiometer. This variable resistor has three terminals. The resistance between the two outside terminals remains constant and the slider moves on the resistance element so the resistance between the center terminal and the other two may be varied. A control of this type may be a wire-wound control or it can also be a carbon control. Most potentiometers are fairly high resistance units and are carbon controls. However, in some applications, for example in color TV receivers, you will run into some very low resistance potentiometers, and these are wire-wound controls. Potentiometers will be found in all types of electronic equipment.

In electronic equipment you will run into all three types of resistors: carbon, wire-wound, and deposited-film. Most often these will be fixed resistors; in other words, they will have a certain value which cannot be changed.

You will also encounter many tapped resistors; this type resistor is always a wire-wound resistor.

In addition, you will run into variable resistors; these may be either wire-wound or carbon resistors. If a variable resistor has two terminals it is called a rheostat; if it has three terminals it is called a potentiometer.

Thus resistors are classified as carbon, wire-wound, and deposited-film; these resistors may be either fixed, tapped, or variable. Variable resistors are divided into rheostats and potentiometers.

Now let's study some of the important uses of resistors.

Using Resistors To Reduce Voltage

We mentioned earlier that in a radio or TV receiver as well as industrial control equipment, many different operating voltages may be required, and these voltages must be obtained from a single power supply to keep down the cost of the equipment. Resistors are often used to reduce the voltage from the power supply to the required value.

SERIES-DROPPING RESISTORS

One of the most common applications of a resistor where it is used to drop voltage is the series-dropping resistor. In this case, the resistor is placed in series with the load, and the current flowing through it produces a voltage drop across the resistor. Small table model radio receivers are an example in which this use of a resistor may be found. In these radio receivers, the heaters of the various tubes are connected in series and they are operated directly from the power line. However, in some cases, the total voltage required by the heaters connected in series may be somewhat less than the power line voltage. In such a case, a resistor is placed in series with the heaters, and part of the voltage is used up by the resistor so that each tube heater gets the correct voltage.

An example of a circuit of this type is shown in Fig. 4. Here we have four tubes with their heaters connected in series. Notice that three of the tubes operate with a heater voltage of 12.6 volts. When the three tubes are connected in series they will require three times this voltage or 37.8 volts. These three tubes are connected in series with the fourth tube that operates with a heater voltage of 50 volts so that the total voltage required by the four tubes in series is 87.8 volts. However, to operate this string of four tubes from a 120-volt power line means that we have to get rid of approximately 32 volts in some

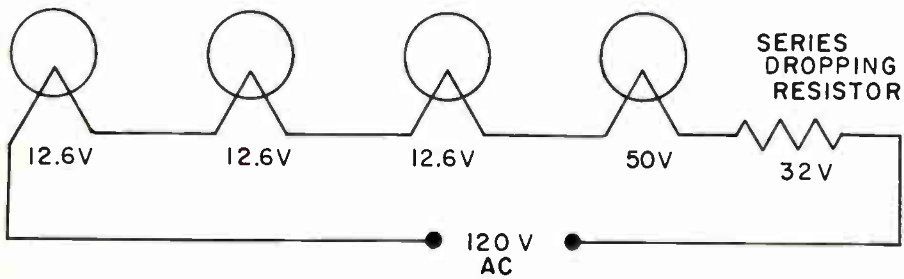


Fig. 4. A series tube heater string, with a series-voltage-dropping resistor.

way in order to get the correct voltage on each tube. Here the series-dropping resistor is connected in series with the tubes. If the tubes operate with a heater current of .15 amps, you can find the value of the resistor required by using Ohm's Law

$$R = \frac{E}{I}$$

and substituting 32 for E and .15 for I, we would find that a resistor with a resistance of 213 ohms would be required. Actually, it probably would be impossible to get a 213-ohm resistor but a 210-ohm or a 215-ohm resistor would certainly be close enough. As a matter of fact, even with a 200-ohm resistor in the circuit, the heater voltage on the tubes would be only slightly over the rated value and should not appreciably affect tube life.

Series-dropping resistors of this type are also to be found in some television receivers. Of course, in the TV receiver there will be far more tubes, and the chances are that the voltage that the resistor must drop will be less than the value in the preceding example, but its purpose is the same, to reduce the line voltage to the value required by the series-heater string.

Series-dropping resistors are also used in dc circuits An example of this type of arrangement is shown in Fig. 5. Here the load, which is represented by R_2, requires an operating voltage of 100 volts. The power supply voltage is 250 volts. The series-dropping resistor R_1 is used to reduce or drop the power supply voltage from 250 volts down to 100 volts for the load. This means that there will be a voltage drop of 150 volts across the series-dropping resistor.

The value of resistance required in R_1 will depend upon the resistance of R_2. Normally the load will have a certain resistance, and when it is operated at the correct voltage a certain current will flow through the resistance. As an example, if the resistance of R_2 is 50,000 ohms, when a voltage of 100 volts is applied across this load, a current of 2 ma will flow through it.

We can use this information to determine the value of the series-voltage-dropping resistor. We know that we must drop a voltage of 150 volts, and since it will be in series with R_2, the current flowing through it will be the same as the current through R_2, in other words 2 ma.

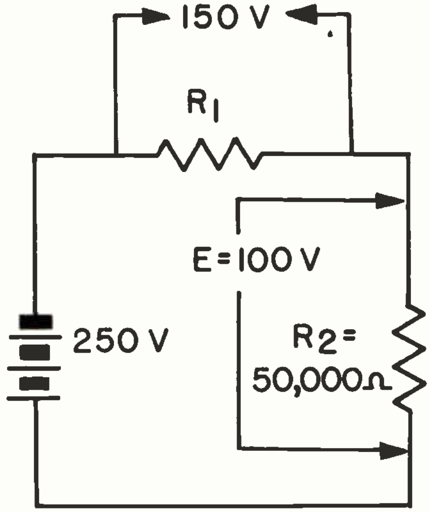


Fig. 5. The series-voltage-dropping resistor is used to reduce the power supply voltage from 250 volts to 100 volts for the load R_2.

We can use this information and Ohm's Law to find the value of R_1.

$$R = \frac{E}{I}$$

$$R = \frac{150}{.002} = 75{,}000 \text{ ohms}$$

In this particular instance we do not even have to resort to Ohm's Law to find the value of R_1. We know the resistance of R_2 is 50,000 ohms and it will have a voltage of 100 volts across it. We want R_1 to have 150 volts across it or 1-1/2 times the voltage across R_2. Therefore the resistance of R_1 must be 1-1/2 times the resistance of R_2. Since R_2 has a resistance of 50,000 ohms, then R_1 must have a resistance of 75,000 ohms.

An arrangement of this type works out very nicely as long as the resistance of R_2 remains constant. However, in many cases in electronic equipment the resistance of the load varies. When this happens, the total resistance in the circuit will vary and this will cause the current to vary. When the current in the circuit varies, the voltage drop across R_1 will vary and when this happens, the voltage across R_2 will also vary. In some cases a fairly wide variation in voltage across R_2 will not present any problem, but there are many applications where we may want to keep the voltage across R_2 reasonably constant. Under these circumstances a simple series-dropping resistor of the type shown in Fig. 5 is not particularly satisfactory.

Let us look at Fig. 6 to see how a variation in the resistance of R_2

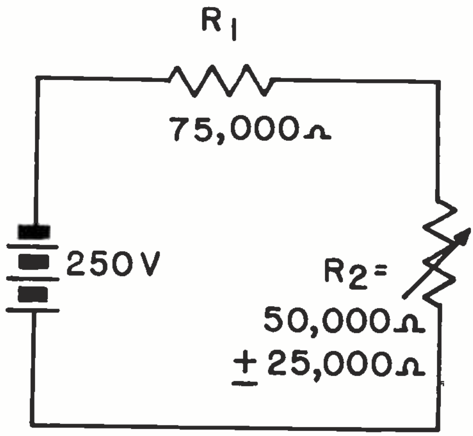


Fig. 6. When the value of R_2 varies, the voltage across it will also vary.

can affect the voltage across it. Notice that here we have represented R_2 as the variable resistance load. Suppose the resistance of R_2 goes down to 25,000 ohms. Now, we have a total resistance in the circuit of 25,000 ohms for R_2 and 75,000 ohms for R_1. This means the total resistance in the circuit will be 100,000 ohms. Using Ohm's Law we find that the current will be

$$I = \frac{E}{R}$$

$$I = \frac{250}{100{,}000} \text{ amps}$$

We can multiply the numerator in this expression by 1,000 and get our answer directly in milliamperes, or perform the division as indicated. If we perform the division as indicated we will find that the current is .0025 amps. If we multiply by 1,000 first, we will get our answer in milliamperes, and in this case

$$I = 2.5 \text{ ma}$$

In either case, we see that the current has increased from 2 ma to 2.5 ma. Now the voltage drop across R_1 will be greater than 150 volts. The actual voltage drop will be 187.5 volts. Since the power supply voltage is 250 volts this leaves us with a voltage of 62.5 volts across the load.

Now let us see what happens when the resistance of R_2 increases. Suppose a resistance of R_2 increases by 25,000 ohms and becomes 75,000 ohms. Now we have R_1 and R_2 in series and since each resistor has a value of 75,000 ohms, half of the voltage will be across each resistor. This means that the voltage across R_2 will increase from 100 volts to 125 volts.

From the preceding we see that if the resistance of R_2 does change, we get quite a substantial voltage variation across the resistor. If the resistance goes down 25,000 ohms, the voltage across the load drops to 62.5 volts, and if the resistance goes up 25,000 ohms, the voltage goes up to 125 volts. A change of this type could appreciably affect the performance of a circuit and in most cases we would have to take steps to prevent such a wide voltage variation.

BLEEDER RESISTORS

Wide voltage variations such as we encountered in the circuit shown in Fig. 6 can be reduced substantially by means of a bleeder resistor. Fig. 7 shows how a bleeder resistor is connected into the circuit. Notice that the bleeder, R_3, is connected in parallel with the load R_2.

The idea in back of the bleeder is to help keep the current flowing through R_1 constant. If the current

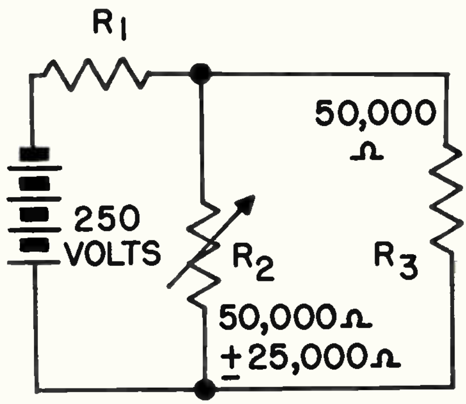


Fig. 7. The bleeder resistor R_3 is connected in parallel with the load R_2 to stabilize the voltage across the load.

flowing through R_1 is constant then the voltage drop across it will be constant and this in turn will mean that the voltage across the load will remain constant.

The higher the bleeder current, the more closely we can keep the voltage constant across the load. This is due to the fact that if the bleeder current is large, it makes up most of the current flowing through R_1, and since it will remain constant, the current changes through R_1 due to variations in the load resistance will be held to a minimum.

Let us look at an example and see how the bleeder actually can help regulate the voltage. Let us take R_2 with a nominal resistance of 50,000 ohms, as before, and connect a bleeder resistor, R_3, in parallel with R_2. Let us select a bleeder that also has a resistance of 50,000 ohms. As before, we want to maintain the voltage across the load constant at 100 volts.

With a voltage of 100 volts across the load resistor, with its value at 50,000 ohms, the current through the load will be 2 ma as before. Since the bleeder is in parallel with the load, 2 ma will also flow through it. This means that the total current in the circuit will be 4 ma.

Now we need to select a series-dropping resistor, R_1. As before, since the power supply voltage is 250 volts, the series-dropping resistor must drop 150 volts. Going back to Ohm's Law, we have

$$R_1 = \frac{150 \text{ volts}}{.004} = 37{,}500 \text{ ohms}$$

Therefore with a series-dropping resistor of 37,500 ohms we will get a voltage drop of 150 volts across R_1, and with a load, R_2 of 50,000 ohms, and a bleeder of 50,000 ohms we will have a voltage of 100 volts across the load and across the bleeder.

With the load and the bleeder in parallel, the total resistance of the two in parallel will be 25,000 ohms. Remember that they are of equal value and therefore the parallel resistance is one-half the resistance of either resistor. Now let's see what happens when the value of R_2 decreases to 25,000 ohms. We will see how this affects the voltage across the load.

When R_2 goes down to 25,000 ohms, we have the 25,000 ohms load in parallel with the 50,000 ohms bleeder. The parallel resistance of the combination is 16,666 ohms which we can round off to 16,500 ohms to simplify our calculations. Total resistance in the circuit will be 16,500 plus the resistance of R_1, which is 37,500 ohms, or a total of 54,000 ohms. The current that will flow in the circuit can be found from Ohm's Law

$$I = \frac{250}{54{,}000} = .0046 \text{ amps}$$

A current of .0046 amps (4.6 ma) flowing through R_1 will produce a voltage drop of

$$E = .0046 \times 37{,}500 = 172.5 \text{ volts}$$

With a voltage drop of 172.5 volts across R_1, the remainder of the voltage will be dropped across R_2 and the bleeder resistor R_3. Subtracting 172.5 from 250 will give us 77.5 volts. Remember that in the preceding example shown in Fig. 6, when the value of R_2 went down to 25,000 ohms, the voltage across it dropped to 62.5 volts. Simply connecting the bleeder resistor with the same value as the nominal value of R_2 in parallel with the load improved the voltage regulation by 15 volts when the resistance of R_2 went down.

When the resistance of R_2 increases by 25,000 ohms, you will have the load R_2, which will then have a resistance of 75,000 ohms, in parallel with the bleeder R_3, which has a resistance of 50,000 ohms. The parallel combination of a 75,000-ohm and a 50,000-ohm resistor will give us a parallel resistance of 30,000 ohms. This resistance in series with R_1 will give us a total resistance of 67,500 ohms in the circuit. With this resistance in the circuit the current flow in the circuit will be .0037 amps and the voltage drop across R_1 138.75 volts. This means that the remainder of the voltage will appear across the load and bleeder and in this case would be 111.25 volts. This compares with the voltage of 125 volts which we found we would get across the load without the bleeder when the resistance of the load increased by 25,000 ohms.

In other words, without a bleeder connected across the load, as the value of the load resistance varied 25,000 ohms above and below its nominal value of 50,000 ohms, its voltage varied from 62.5 volts to 125 volts. With a 50,000-ohm bleeder connected across the load, when the load resistance varied 25,000 ohms above and below its nominal value, the voltage across the load varied from 77.5 volts to 111.25 volts. You will notice that connecting the bleeder in parallel with the load has had a substantial effect in maintaining a more constant voltage across the load.

The larger the bleeder current, the better the voltage regulation will be. In fact, if the nominal value of R_2 is 50,000 ohms and we put a 5,000-ohm bleeder in parallel with it we get very little voltage change at all as the resistance of R_2 varies between 25,000 and 75,000 ohms. With a bleeder of this size, the total series current flow would be about .022 amps. This would require a series-dropping resistor of approximately 6800 ohms. When the resistance of R_2 dropped from 50,000 ohms to 25,000 ohms, the current in the circuit would increase from .022 amps to slightly less than .023 amps. The increase in the voltage drop across R_1 would be only 4 volts, which means that the voltage drop across the load would change from 100 volts to 96 volts. Similarly, if the resistance of R_2 increased from 50,000 ohms to 75,000 ohms, the current in the circuit would drop only about one half a milliampere and the voltage drop across R_1 would change from 150 volts to approximately 147 volts. Therefore the voltage across the load would increase to 103 volts. As you can see, with a low-resistance bleeder that draws a high bleeder current, the voltage across the load remains almost constant, even with the same variation in the load that produced a wide voltage variation before.

Since a bleeder that draws a large current regulates the voltage so well you might wonder what the problem is. In a circuit where we want to keep the voltage across the load constant, why not simply put a bleeder across the load that will draw a high current and maintain the voltage constant? The answer is that the bleeder current serves no purpose other than to regulate the voltage across the load. In so far as performing any other function is concerned, it is wasted. There is a limit to how much current we can take from a power supply. The more bleeder current we draw the bigger and more costly the power supply must be. Therefore in applications where a bleeder is used, in most cases some compromise is reached and a value of bleeder is selected that will give the voltage regulation required, and no more. Of course in some applications where very precise regulation is required, we have to waste the power in the bleeder in order to get this regulation. In this case, we must build a power supply capable of supplying the required current even though it may be quite costly.

Series-dropping resistors and bleeders are often used in tube and transistor circuits. You have already studied the triode tube and will remember that it has three elements, a cathode, a grid and a plate. The pentode tube is a tube with five elements. It has the same three elements as the triode tube plus a suppressor grid and a screen grid. The suppressor grid is usually connected to B- or to the cathode. The screen grid is connected to B+ and usually operated at a voltage somewhat less than the plate voltage. The voltage required for the screen is obtained through a series-dropping resistor from the same power supply that supplies the plate voltage. If the screen voltage must be maintained constant then you will often find a bleeder in the screen circuit for this purpose.

SUMMARY

Now let us review what you have studied in this section on resistors and how they are used to reduce voltage. Resistors are sometimes used in the heater circuit of small radios or in the heater circuit of television receivers to reduce the line voltage to the value required by series-connected tube heaters. Resistors are also used to reduce the voltage to a load so that the load can be operated from a power supply that has a somewhat higher voltage than the voltage required by the load.

Resistors are used as bleeders to stabilize or regulate the voltage across a load. The bleeder consumes or wastes a certain amount of current, but this extra current that flows through the bleeder remains constant and helps maintain the voltage drop across the series-dropping resistor constant. The larger the current consumed by the bleeder in comparison to the current consumed by the load, the better the regulation across the load will be. However, since the bleeder current is waste current and performs no useful purpose other than to regulate the voltage across the load, we normally do not use any more bleeder current than is necessary to get the degree of regulation that is required across the load. We keep the bleeder current as low as possible in order to keep the cost and size of the power supply as low as feasible. However, in some applications where very precise regulation is required, a large bleeder current is used and we simply have to go to the expense of making the power supply as large as necessary to supply this current along with the load current.

Now to help you be sure that you understand this section of the lesson and to review it, answer the following self-test questions.

SELF-TEST QUESTIONS

(a) What is the purpose of a series-dropping resistor?

(b) If two tubes that each require a heater voltage of 35 volts are connected in series with two additional tubes that require a heater voltage of 6 volts each and the four heaters connected in series are to be operated from a 120-volt power line, how much voltage must the series-dropping resistor drop in order to provide the correct heater voltages for the tubes?

(c) In the preceding example, if the current drawn by the tubes is .3 amperes, what would be the value of the series-dropping resistor to use?

(d) What is the purpose of a bleeder?

(e) Which type of bleeder is more effective, a high-resistance bleeder that draws very little current, or a low-resistance bleeder that draws a substantial current?

(f) What consideration other than regulation must be kept in mind in selecting a bleeder?

(g) If a load, which has a resistance of 20K ohms and requires an operating voltage of 200 volts, is connected across a power supply that has an output voltage of 300 volts, what should the resistance of the series-voltage-dropping resistor be?

Power In Electrical Circuits

We have already mentioned several times that the unit of electrical power is the watt. The watt tells us how much electrical energy or power is being expended or used in a circuit. You have probably run into this unit many times, and you are no doubt familiar with different size light bulbs - they are rated in watts. A 60-watt bulb consumes 60 watts of electrical energy when it is lit. A 100-watt electric bulb consumes 100 watts of electrical energy - almost twice as much power as a 60-watt bulb.

Now let us go ahead and learn more about exactly how much electrical energy the watt represents.

THE WATT

The power in an electrical circuit is equal to the product of the voltage times the current. Expressing this as a formula we have

$$P = E \times I$$

We generally drop the times sign and simply write the formula as

$$P = EI$$

This formula tells us that if a generator is supplying a voltage of 1 volt to a circuit and the current flowing in the circuit is 1 amp, the power being supplied by the generator is 1 watt. If the generator is supplying a voltage of 10 volts and a current of 2 amps, the power being supplied by the generator is $10 \times 2 = 20$ watts.

From the expression for power we see that the current used by a 100-watt electric light bulb will be slightly less than 1 amp when the bulb is operated from a 120-volt power line. Since the product of the voltage times the current equals 100 watts, we can find the current by rearranging the power formula to

$$I = \frac{P}{E}$$

and substituting 100 watts for P and 120 volts for E

$$I = \frac{100}{120} = .83 \text{ amps}$$

A 60-watt bulb will draw somewhat less current,

$$I = \frac{60}{120} = .5 \text{ amps}$$

The power in an electrical circuit can also be expressed in terms of voltage and resistance or in terms of current and resistance by combining the power formula with Ohm's Law. For example from Ohm's Law we know that

$$E = I \times R$$

If we substitute $I \times R$ for E in the power formula we have

$$P = I \times R \times I = I \times I \times R$$

If we drop the times sign and write $I \times I$ as I^2 (this is called I squared) we will have the formula

$$P = I^2R$$

You will see this expression many times in electronics. Remember that the term I^2 means $I \times I$. This form of the power equation is often used to determine the wattage rating of a resistor in a circuit where the current through the resistor and the resistance of the resistor are known.

Going back to our original power formula

$$P = E \times I$$

and Ohm's Law in the form

$$I = \frac{E}{R}$$

if we substitute $\frac{E}{R}$ for I we will get

$$P = E \times \frac{E}{R}$$

which we usually write as

$$P = \frac{E^2}{R}$$

where the expression E^2 means $E \times E$. This form of the power equation can be used where we know the voltage across a part and the resistance of the part or circuit.

These three forms of the power formula are extremely important; you will have many occasions to use them. Therefore you should take the time now to memorize them to save having to look them up in the future. Even though you may not have to do a great deal of calculating involving these formulas you should know what the formulas are so you will be able to approximate the power used in a circuit. The three formulas are:

$$P = EI$$

$$P = I^2R$$

$$P = \frac{E^2}{R}$$

WATTAGE RATING OF RESISTORS

Earlier we mentioned that resistors were made with different wattage ratings. When current flows through a resistor a certain amount of power is used, or as we say dissipated. The wattage rating of a resistor tells us how much power the resistor can dissipate. If the power being dissipated by the resistor is greater than its wattage rating, the resistor will soon burn out.

We mentioned previously that carbon resistors are made in 1/2-watt, 1-watt and 2-watt sizes. Also, you will occasionally run into very small carbon resistors that are rated at 1/3-watt.

When we say that a resistor is a 1-watt resistor we mean that the power it can dissipate or handle is 1-watt. This means that the product of the current squared through the resistor times the resistance of the resistor must be equal to 1 watt or less.

Actually, in the case of carbon resistors, it is not a good idea to use them at their maximum rating. Carbon resistors have a tendency to change value if they get too hot. A 1-watt carbon resistor that is actually dissipating one full watt of electric power will get quite warm after it has been in operation for some time. Eventually, this will cause the resistor to change value. There is no way of predicting how much the resistor will change value - the change might be slight and may not affect the performance of the circuit. On the other hand, the change might be large enough to appreciably affect the performance of the circuit. The usual practice is to use resistors having a wattage rating almost double that actually required. In other words, if the power that a resistor must dissipate is 1 watt or slightly over 1 watt then you would use a 2-watt resistor. If the resistor must dissipate approximately 1/2 watt, it is best to use a 1-watt resistor to avoid any possible problem in the future due to the resistor overheating and changing value.

The deposited film type of resistor is made in ratings of 2 watts up to about 10 watts. These resistors are somewhat larger than carbon resistors and are able to handle the larger power. They do not have the tendency to change value that carbon resistors have and therefore can be operated closer to their full wattage rating. However, even in the case of deposited film resistors, most manufacturers usually allow a reasonable safety factor to prevent the resistors burning out after they have been in service for some time. As an example, if the wattage being dissipated by a deposited film resistor is almost four watts, a 5-watt resistor is usually used to provide some safety factor.

There is almost no limit to the wattage ratings in which wire-wound resistors can be made. In radio and television receivers you will seldom find wire-wound resistors rated at higher than 50 watts. However, in radio and TV transmitting equipment and in industrial control equipment you may find wire-wound resistors that are capable of dissipating several thousand watts or more. These high-wattage resistors are made with a large-size resistance wire that can carry the high current and are wound on a large diameter tube in order to provide room for the wire needed to get the required resistance.

The same safety rule for carbon resistors is generally followed in the case of wire-wound resistors. Usually a resistor having approximately twice the wattage rating that the resistor must dissipate is used. Manufacturers follow this practice because the larger wattage resistor will be a larger physical size and better able to get rid of the heat that is produced by the electrical energy that the resistor must dissipate and also because it provides a safety factor - there is far less chance of the resistance wire burning out if the overrated resistor is used.

TRANSFERRING POWER

A common problem in electronics is transferring power from one circuit to another or from one device to another. As an example, consider the output tube or transistor in a radio receiver and the loudspeaker in the receiver. The output tube or transistor develops audio power; in other words it develops sound power in electrical form. The problem is to get the maximum amount of power over to the speaker from the output tube or transistor. Now let us consider what is involved in this power transfer.

The tube or transistor acts as the power source. The speaker acts as a load. Thus we have a situation somewhat similar to a battery with a load connected across it. We have seen this circuit many times.

However, there is one thing we have omitted in the circuits we have looked at before. All electrical parts have resistance. This is true of transistors, tubes, coils, transformers, generators and batteries. The usual practice is to simply represent a battery by the battery symbol. However, to be precise, we should represent it by the battery symbol and a resistor in series with it to represent the internal resistance of the battery. By internal resistance of the battery we mean the opposition that current encounters in flowing from the positive terminal of the battery through the battery to the negative terminal.

In Fig. 8 we have shown a circuit representing a battery with internal resistance and a load connected across the battery. The battery we have shown is a 10-volt battery with an internal resistance of 5 ohms. Thus in any circuit we consider, the total resistance in the circuit will be the resistance of R_1 plus the internal resistance of the battery, which is 5 ohms. This total resistance is the opposition to current flow in the circuit, and is the value we must use in performing any calculations to find the current flowing in the circuit or the power dissipated in the circuit.

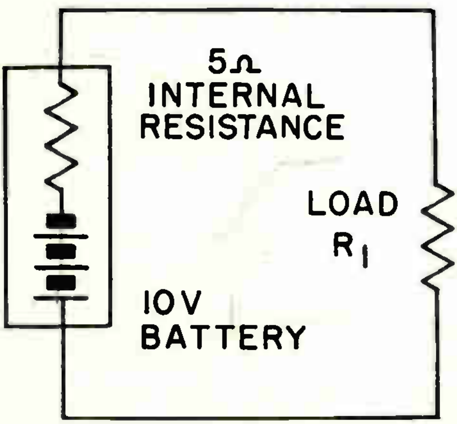


Fig. 8. Circuit representing a battery with internal resistance and a load across the battery.

The fact that the battery has internal resistance brings about several interesting things. First, let us assume that the resistance of R_1 in the circuit is 20 ohms; this means that the total resistance in the circuit will be 25 ohms. Since the battery potential is 10 volts this means that the current flowing in the circuit will be

$$I = \frac{10}{25} = .4 \text{ amps}$$

The .4 amps flowing through the battery resistance will produce a voltage drop across this resistance. This voltage drop will be

$$E = .4 \times 5 = 2 \text{ volts}$$

Thus the actual voltage available at the battery terminals will be only 8 volts. This is quite common in batteries and generators; part of the voltage produced by the battery or generator is lost due to the internal resistance of the device.

Now let us find how much power is actually being supplied by the battery, and how much is being received by the load. The power supplied by the battery is equal to the battery voltage times the current or $P = 10 \times .4 = 4$ watts.

The power being dissipated by the load can best be found by the power formula

$$P = I^2R$$
$$P = .4 \times .4 \times 20 = 3.2 \text{ watts}$$

Now let us replace the 20-ohm resistor that we have in the circuit for R_1 with a 10-ohm resistor and see what happens in the circuit. With a 10-ohm resistor in the circuit the total resistance in the circuit will be 10 ohms plus the 5 ohms internal resistance of the battery or a total of 15 ohms. The value of the current in the circuit will be

$$I = \frac{10}{15} = .67 \text{ amps}$$

Now the power being supplied by the battery will be

$$P = 10 \times .67 = 6.7 \text{ watts}$$

and the power dissipated by the resistor R_1 will be

$$P = .67 \times .67 \times 10 = 4.5 \text{ watts (approximately)}$$

If we replace R_1 with a 6-ohm resistor, we will have a total resistance of 11 ohms in the circuit. This will give us a current flow of .9 amps and the power dissipated in R_1 will be 4.8 watts.

If we substitute a 5-ohm resistor for R_1, the total resistance in the circuit will be 10 ohms and the current will be 1 amp. The power dissipated by R_1 will then be 5 watts. The total power generated by the battery will be 10 watts.

If we substitute a 4-ohm resistor for R_1 the total resistance in the circuit will be 9 ohms and the current flow 1.1 amps. The power dissipated by R_1 will then be 4.8 watts. If we continue to reduce the size of the resistance we will find that the power dissipated by the resistor will continue to go down.

Notice that as we started with a 20-ohm resistor and reduced the size of it to 5 ohms, that the power dissipated by the resistor increased until it reached a maximum value at 5 watts. When we reduced the resistance of the resistor below 5 ohms, the power starts to decrease.

The significant thing to notice here is that we obtain maximum power in the resistor when the resistance of the resistor is equal to the internal resistance of the battery.

This situation will always exist. We will get maximum power in the load when the load resistance is equal to the generator resistance. However, this condition may not always be desirable because the total power being generated by the generator when we have a 5-ohm resistance in the circuit is 10 watts. Since we are dissipating only 5 watts in the resistor, half of the power is being wasted in the battery itself. Thus, when maximum power transfer is being obtained, the over-all efficiency of the system is only 50% - half the power being produced by the source is transferred to the load. With a higher load resistance, the efficiency improves and under some circumstances we may be willing to get something less than full power transfer to get better efficiency. For example, when the value of R_1 was 20 ohms we had a power of 3.2 watts in the resistor and the total power produced by the generator was 4 watts. This is an efficiency of 80%; in other words, 80% of the power produced by the battery is transferred to the load. Even though this is less than the power we get when the load is matched or equal to the battery resistance, the efficiency of the power transfer is much better.

The important point for you to remember from this section of the lesson is that we will get maximum power transfer from a voltage source to a load when the resistance of the load is equal to the internal resistance of the source. Under these conditions we say that the load is matched to the generator. Under these conditions, the efficiency of the power transfer will be 50% - half the power will be dissipated in the load and the other half lost in the generator or the voltage source. Under some circumstances it is better to get somewhat less power transferred from the source to the load in order to get better efficiency.

SUMMARY

This section of this lesson is an extremely important one, and you should be sure that you understand it completely before going on.

You learned that the unit of power is the watt. The power in a circuit can be obtained from any one of the three formulas

$$P = EI$$

$$P = I^2R$$

$$P = \frac{E^2}{R}$$

You should memorize these three formulas; you will need them over and over again in your career in electronics.

Remember that resistors are made in different wattage ratings. Remember too, that carbon resistors have a tendency to change value if they are operated near their maximum wattage rating. If you have to replace a carbon resistor in a piece of electronic equipment you can use the same wattage rating resistor as used by the manufacturer or one having a higher wattage rating, if there is room for the larger resistor in the circuit. The usual safety factor allowed in electronic equipment is to use a resistor having twice the wattage rating that it must dissipate. This will reduce the possibility of resistor failure after the equipment has been in use for some time.

Remember that maximum transfer of power from a generator to a load can be obtained when the resistance of the load matches the resistance of the generator. When maximum power transfer is obtained the efficiency is only 50%; this means that only half the power produced by the generator reaches the load. Under some circumstances we will be satisfied with a somewhat lower power in the load in order to obtain better efficiency.

Now do the self-test questions on this section carefully. If you have difficulty with any of the questions it is a sign that you need to spend more time on this important section.

SELF-TEST QUESTIONS

(h) What is the unit of electrical power or energy?

(i) If a battery has a voltage of

15 volts, and it is supplying a current of 3 amps, what is the power being supplied by the battery?

(j) What do we mean when we say that a resistor is dissipating 10 watts?

(k) If the voltage across a 1000-ohm resistor is 100 volts, how much power is the resistor dissipating?

(l) If the voltage across a 5000-ohm resistor is 50 volts, what is the power dissipated by the resistor?

(m) If the current through a resistor is 2 amps, and the resistance of the resistor is 25 ohms, what is the power dissipated by the resistor?

(n) If the current through a 100K resistor is 10 ma, how much power is the resistor dissipating?

(o) A battery has a voltage of 9 volts and an internal resistance of 3 ohms. What size resistor should be used as a load in order to get maximum power transfer?

(p) A 20-volt battery has an internal resistance of 10 ohms. What size resistor should be be connected across it as a load in order to get maximum power transfer? How much power will be produced by the battery and how much power will be transferred by the load?

Resistor Values

We mentioned earlier that there are more resistors used in electronic equipment than any other parts. A small radio receiver generally has somewhere between ten and fifteen resistors. A black and white TV receiver usually has approximately 50 resistors and a color TV receiver may have twice that many. Usually you can figure that a piece of electronic equipment will have somewhere between three and five resistors at least, for each tube or transistor used in the equipment.

We have already gone into the methods used to indicate the value of a resistor. You know that the unit of resistance is the ohm. If the voltage across a resistor is 1 volt and a current of 1 amp flows through the resistor, the resistance of the resistor is 1 ohm. You will remember this important relationship is expressed in Ohm's Law.

You are also familiar with the fact that large values of resistors are found in electronic equipment and that we use the symbol K to represent 1000 and the letter M for megohms which represents 1,000,000 ohms. Thus a 2,200-ohm resistor will often be marked 2.2K on a diagram and a 4,700,000-ohm resistor will often be marked 4.7M or 4.7 megs.

Since there are 1000 ohms in 1K-ohm and 1,000,000 ohms in a megohm, it follows that there are 1000K-ohms in a megohm. You need to be able to convert between K-ohms and megohms because on some diagrams a resistor having a resistance of 100K-ohms may be marked as .1 meg, and you must realize that both represent the same value. To convert from ohms to K-ohms you move the decimal point three places to the left, and to convert from K-ohms to ohms you move it three places to the right. To convert from ohms to megohms you move the decimal point six places to the left and to convert from megohms to ohms you move it six places to the right. To convert from K-ohms to megohms you move the decimal point three places to the left and to convert from megohms to K-ohms you move it three places to the right. Remember that to convert from the smaller unit to the larger unit you move the decimal point to the left, and to convert from the larger unit to the smaller unit you move the decimal point to the right.

TOLERANCES

The most frequently encountered resistor is the molded-carbon resistor. The molded-carbon resistor is made with three different tolerance values. The resistor is made with a 5% tolerance, a 10% tolerance or a 20% tolerance. The tolerance indicates how much the resistor may vary in resistance from its indicated value. For example, a 100-ohm resistor with a 5% tolerance will be within 5% of 100 ohms. This means that its value might be 5% below or as much as 5% above 100 ohms. 5% of 100 is 5 ohms and therefore the resistance of the resistor may be any value between 95 and 105 ohms.

In this case of a 10% resistor, the value may vary 10% above or below 100 ohms. 10% of 100 is 10 ohms and therefore the resistance of this resistor can be any value between 90 ohms and 110 ohms. A 20% resistor can have a tolerance of 20 ohms and therefore the resistance could be any value between 80 ohms and 120 ohms.

The closer the tolerance of a resistor, the more expensive the resistor is. Thus a 5% resistor is more expensive than a 10% resistor and a 10% resistor is more expensive than a 20% resistor. However, with today's modern automatic resistor making machinery, there is very little difference between the price of a 10% resistor and a 20% resistor and therefore you don't run into too many 20% resistors in electronic equipment any more. There is considerable difference between the price of a 5% resistor and a 10% resistor, and therefore most manufacturers will use 10% resistors wherever they can. The 5% resistors are found only in the more critical circuits where it is important that the value of the resistor be held to a close tolerance.

EIA VALUES

The EIA (Electronic Industries Association) has set up standard carbon resistor values. You cannot buy any value carbon resistor you might want - you have to buy one of the standard values. The standard values are arranged so that you can get a resistor within 5% of any required value. For example, if you determined that in a certain circuit you needed a 53,000-ohm resistor, you cannot buy a 53,000-ohm carbon resistor - they are not made in this size. However, you can buy a 51,000-ohm resistor and a 56,000-ohm re-

Ohms	Ohms	Ohms	Ohms	Ohms	Ohms	Ohms	Ohms	Ohms	Ohms	Megs	Megs
0.24	1.1	5.1	24	110	510	2400	11K	51K	240K	1.1	5.1
0.27	**1.2**	**5.6**	**27**	**120**	**560**	**2700**	**12K**	**56K**	**270K**	**1.2**	**5.6**
0.30	1.3	6.2	30	130	620	3000	13K	62K	300K	1.3	6.2
0.33	**1.5**	**6.8**	**33**	**150**	**680**	**3300**	**15K**	**68K**	**330K**	**1.5**	**6.8**
0.36	1.6	7.5	36	160	750	3600	16K	75K	360K	1.6	7.5
0.39	**1.8**	**8.2**	**39**	**180**	**820**	**3900**	**18K**	**82K**	**390K**	**1.8**	**8.2**
0.43	2.0	9.1	43	200	910	4300	20K	91K	430K	2.0	9.1
0.47	**2.2**	**10**	**47**	**220**	**1000**	**4700**	**22K**	**100K**	**470K**	**2.2**	**10**
0.51	2.4	11	51	240	1100	5100	24K	110K	510K	2.4	11
0.56	**2.7**	**12**	**56**	**270**	**1200**	**5600**	**27K**	**120K**	**560K**	**2.7**	**12**
0.62	3.0	13	62	300	1300	6200	30K	130K	620K	3.0	13
0.68	**3.3**	**15**	**68**	**330**	**1500**	**6800**	**33K**	**150K**	**680K**	**3.3**	**15**
0.75	3.6	16	75	360	1600	7500	36K	160K	750K	3.6	16
0.82	**3.9**	**18**	**82**	**390**	**1800**	**8200**	**39K**	**180K**	**820K**	**3.9**	**18**
0.91	4.3	20	91	430	2000	9100	43K	200K	910K	4.3	20
1.0	**4.7**	**22**	**100**	**470**	**2200**	**10K**	**47K**	**220K**	**1 meg**	**4.7**	**22**

Fig. 9. Standard EIA carbon resistor values.

sistor and therefore you would select one of these two values - the exact one that you would select would depend upon whether you want the resistance to be a little higher than the calculated value or a little lower.

Standard EIA carbon resistor values are shown in Fig. 9. All are available with a 5% tolerance; the values in bold type are also available in 10% tolerances. You normally cannot buy 20% resistors for replacement purposes, but either 5% or 10% resistors are satisfactory replacements.

COLOR CODE

Carbon resistors are identified by means of a color code. There will be three or four colored bands on a carbon resistor. The resistors with only three color bands are 20% tolerance resistors. The resistors with the four color bands will be either 5% or 10% resistors depending upon the color of the tolerance band. If the fourth band is a gold band the resistor has a tolerance of 5%, and if it is a silver band the resistor has a tolerance of 10%.

The color bands are placed on the body of the resistor nearer to one end than the other as shown in Fig. 10. To read the color band, hold the resistor as shown in Fig. 10. The fourth band, if there is a fourth band, will be either gold or silver and it will be on the right and tell you the tolerance of the resistor. To read the value of the resis-

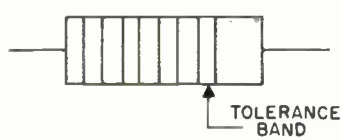


Fig. 10. Resistor values and tolerance are identified by color bands as shown.

Color	1st Figure	2nd Figure	No. of Zeros
Silver			.01
Gold			.1
Black	0	0	none
Brown	1	1	0
Red	2	2	00
Orange	3	3	000
Yellow	4	4	0000
Green	5	5	00000
Blue	6	6	000000
Purple	7	7	
Gray	8	8	
White	9	9	

Fig. 11. Standard resistor color code.

tor start at the left end. The first color band gives you the first significant figure of the resistance value. The second color band gives you the second figure and the third tells you how many zeros to add to get the resistance of the resistor.

Values assigned the various colors are shown in Fig. 11. For example, a resistor with (left to right) red, red, black, and gold bands would have a resistance of 22 ohms. The first and second bands each indicate 2; the black band indicates no zeros. The gold band indicates 5% tolerance. If the resistor were colored orange, orange, red, the first and second bands would each indicate 3, and the red band two zeros, so the value would be 3300 ohms or 3.3K. If the fourth band is silver, the tolerance is 10%, if gold 5%. If there is no fourth band, the tolerance is 20%.

If the third color band on a resistor is gold, it indicates that you multiply the first two numbers by .1 to get the resistor value. Thus a resistor coded red, red, gold, gold is $22 \times .1 = 2.2$ ohms, 5%. If the third color band is silver, you multiply by .01. A resistor coded red, red, silver, gold is $22 \times .01 = .22$ ohms, 5%.

When you start to work on your experimental kits, you will have to use the color code to identify the various resistors in the kit. It is worthwhile to memorize the color code. However, do not spend a great deal of time trying to do it all at once. Look over the color code shown in Fig. 11 two or three times and learn at least what a few of the colors represent. Then after you have finished this lesson go back and take a look at it again and read it through two or three times. If you will do this several times, this, and working with your experimental kits, will soon teach you the code so that you will know it by heart. In your early experimental kits we will give you the resistance value and the color code to help you learn the code and learn to identify the resistors, but you must learn to do this yourself so you will be able to identify resistors in any electronic equipment you may be called upon to maintain or service.

DEPOSITED CARBON RESISTORS

There is another type of carbon resistor known as the deposited carbon resistor. The carbon resistors that you will run into most frequently in electronic equipment can easily be identified because they are color coded by bands as shown in Fig. 10. The deposited carbon resistors are not made in the same way as these carbon resistors. They are made by depositing a layer of carbon on a form. The layer can be controlled closely and can be varied as necessary in order to make resistors that can be held to a very close tolerance. These resistors are usually made in tolerances of 1% and 1/2%. It is not likely that you will run into this type of resistor in commercial entertainment type equipment such as radio or TV receivers, but you will run into them in meters and test equipment where the accuracy of a meter reading will depend upon the accuracy of the resistors in the equipment. This type of resistor is identified by stamping the value of the resistor on the body of the resistor along with its tolerance. If you should have to repair a meter or some other test instrument and replace one of these resistors, it is important that the replacement have exactly the same resistance and tolerance as the original. Of course, if you have to replace a 100K-ohm, 1% resistor and can't get a 1% resistor, you can use a 1/2% tolerance resistor if one is available. 1/2% and 1% resistors are not nearly as readily available from radio and TV parts wholesalers as are the standard 5% and 10% molded carbon resistors. Often these precision resistors, as they are called, must be obtained by ordering them specially through your parts wholesaler.

SUMMARY

You have reached a point in your course now where you should be able to convert from ohms to K-ohms and megohms or from the larger units back to ohms without any trouble. The reason why it is important that you can convert from one unit to the other is that all three units are used by manufacturers on circuit diagrams and you have to know what is meant when any one of the units is used.

It is important for you to remember that resistors are made in only certain standard sizes and that all molded carbon resistors have certain tolerances. The standard tolerances are 5%, 10% and 20%. You will find mostly 10% resistors used in radio and TV receivers; both 10% and 5% resistors will be found in industrial control equipment.

You should start to learn the resistor color code by memory. You will have to use it over and over again to identify resistors in electronic equipment. However, as we point out don't try to memorize it all at once. Read it through two or three times whenever you think about it and you will soon have it memorized.

Remember that 1% and 1/2% resistors are used in test equipment and if you should have to replace any of these resistors be sure that you use an exact value replacement and a replacement having a tolerance of at least as close as the tolerance of the original.

Now answer the following self-test questions on this section of the lesson.

SELF-TEST QUESTIONS

(q) What is 4.7K-ohms equal to in ohms?

(r) Express .39 megohms in K-ohms and in ohms.

(s) What is 680,000-ohms in K-ohms and in megohms?

(t) A 2200-ohm, 10% resistor actually has a value of 2000 ohms. Is this resistor within its rated tolerance?

(u) A 10,000-ohm resistor has a tolerance of 5%. What is the maximum value resistance that the resistor might actually have and still be within tolerance?

(v) Reading from left to right the color bands on a resistor are orange, white, yellow and gold. What is the value and tolerance of the resistor?

(w) If a resistor is color coded brown, black, green and silver, what is its value and tolerance?

(x) If a resistor is color coded green, blue, orange and gold, what is its value and tolerance?

(y) A resistor is colored red, red, red and silver. What is its value and tolerance?

(z) In a certain piece of test equipment a 100K, 1% resistor has burned out. You have available a 100K, 1/2% resistor. Can you use this resistor as a replacement?

Resistors With Special Characteristics

We have already mentioned one of the important characteristics of carbon resistors, that is the fact that they change resistance with changes in temperature. Carbon itself has what we call a negative temperature coefficient. Most elements that carry electricity have a positive temperature characteristic. This means that if a copper wire is carrying a current and the temperature of the wire increases the resistance will increase - the wire has a positive temperature coefficient. Carbon, as an element, does the opposite - if it is heated, its resistance goes down; it has a negative temperature coefficient.

Most carbon resistors have a negative temperature coefficient. However, sometimes due to the way the resistor is made and due to the type of material used as a binder to hold the resistor together, the resistor might actually have a small positive temperature coefficient. What happens is that the carbon in the resistor has a negative temperature coefficient which tends to cause a resistance to go down with an increase in temperature but the other material has a positive temperature coefficient greater than the carbon temperature coefficient and this causes the resistance to go up when the temperature increases. With the two fighting each other, it is hard to predict what will happen. Some resistors will go down slightly in resistance when they are heated, others will increase slightly in resistance.

Wire-wound resistors and deposited-film resistors in general have a positive temperature coefficient. Their resistance will increase as the temperature is increased.

Actually, as far as a general purpose resistor is concerned, we would prefer to have them with a zero temperature coefficient. In other words, we would like their resistance to remain constant as their temperature changes. Changes in resistance due to temperature changes often produce undesirable results in electronic equipment. However, with modern manufacturing techniques, most carbon, deposited-film and wire-wound resistors have such a low temperature coefficient that any change in their temperature in normal operation will not cause their resistance to change sufficiently to cause any serious problem.

In some special applications, however, it is desirable to have a resistor whose resistance value will change with temperature. In this section of the lesson we are going to briefly study two of these devices.

THERMISTORS

A thermistor is a type of resistor that is made of a material whose resistance value varies with changes in temperature. The material used is a form of a semi-conductor material quite similar to the material used in the manufacture of transistors.

Thermistors have a negative temperature coefficient. This means that as the temperature increases, the resistance of the thermistor decreases. The amount that the resistance of the thermistor will change with a given change in temperature depends upon the type of material from which the thermistor is made.

Thermistors are quite often used in circuits where there is liable to be a large current surge when the equipment is first turned on. An example of this type of application is in the power supply of a television receiver that is designed for operation without a power transformer. When the equipment is first turned on, a very high surge current will flow for a short time. A thermistor placed in the circuit can limit this current to a safe value. As the equipment starts to operate, the current flowing through the thermistor causes it to heat so that its resistance value drops very rapidly and by the time the equipment has reached operating temperature and is ready to operate, the resistance of the thermistor has dropped to such a low value that it has very little effect upon the performance of the equipment. Thermistors of this type may have a resistance of 100 ohms or more when they are cold, and a resistance of only a few

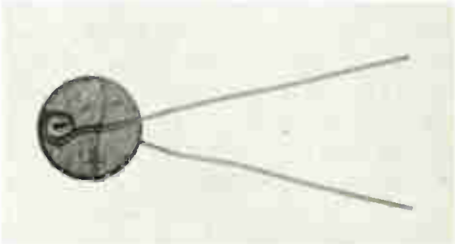


Fig. 12. A typical thermistor, such as might be found in a TV receiver, and its schematic symbol.

ohms when they have reached their normal operating temperature in the receiver.

A photograph of a thermistor such as you might find in the application we mentioned in the television receiver is shown in Fig. 12. Notice that the thermistor is made in the form of a round disc with leads attached to each side of it. The schematic symbol used to represent the thermistor is also shown in Fig. 12.

The thermistor type shown in Fig. 12 is the type you are most likely to encounter in TV receivers. However, there are many other different types. A photo of a number of different types is shown in Fig. 13.

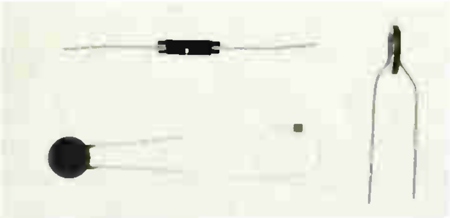


Fig. 13. Thermistors are made in many different shapes as shown above.

These types are often found in industrial electronic equipment. Sometimes they are used to control current surges, in other applications they are used to measure temperatures. You can tell the temperature of the thermistor by measuring its resistance and thus the thermistor can be used for temperature measuring applications.

VARISTORS

Another special type of resistor is the varistor. This resistor is also called a voltage-dependent resistor. This means that the resistance of the device depends upon the voltage

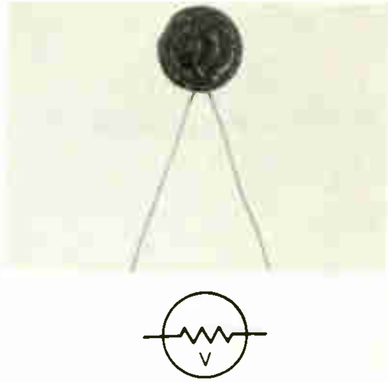


Fig. 14. A typical varistor and its schematic symbol.

across it. In other words, as the voltage across the varistor increases, the resistance of the varistor decreases.

Varistors are used in circuits to protect components from damaging high voltage transients. This could cause component failure. With a varistor in the circuit, as the voltage rises the resistance of the varistor decreases, drawing a large current from the voltage source which will lower the voltage.

Varistors will be found in some color TV receivers. A photo of a varistor is shown in Fig. 14 along with the schematic symbol used to identify it. You will notice that the varistor looks very much like a thermistor - as a matter of fact, it is sometimes difficult to tell them apart from their appearance alone.

HIGH-VOLTAGE RESISTORS

In some applications resistors will be used across circuits where there is a comparatively high voltage. Examples of this type of resistor will be found in many television receivers.

High-voltage resistors must be made quite long in order to keep the voltage from jumping across the resistor. Usually the resistance element is placed on a form in the shape of a spiral curve. An example of this type of resistor is shown in Fig. 15A. The resistor is about 2-1/2 inches long and is used in a circuit in a color TV receiver where the operating voltage is in excess of 6000 volts. The resistor is made by putting the carbon on a form shaped like a spiral that winds around the resistor from one end to the other as shown in the drawing in Fig. 15B.

Resistors of this type are made with a spiral type element because they normally have a very high resistance and by spiralling the resistance element around the form in this way it is possible to get a much longer path, and therefore the high resistance needed can be obtained in a reasonable size. The unit shown in the photograph has a resistance of 66 megohms. It would be difficult to get this much resistance in a resistor of this size and type by any method other than the spiral-wound method. If you should have occasion to replace a resistor of this type in a television receiver or in any other device where high voltages are used, you must be sure to use this special type of high-voltage resistor.

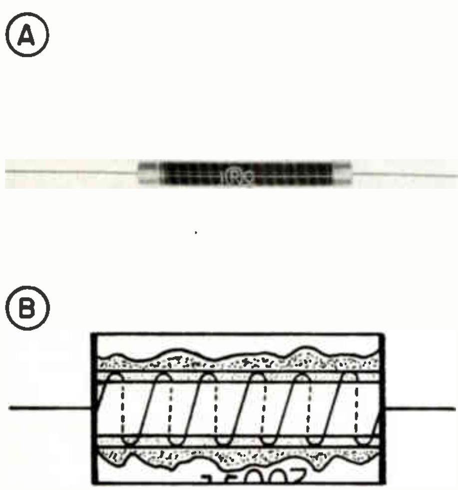


Fig. 15. A high-voltage resistor.

SUMMARY

In this section of the lesson we have briefly studied three special types of resistors - the thermistor, the varistor or voltage-dependent resistor, and the high-voltage resistor. A few years ago these three types of resistors were unknown to the electronics technician, but modern technology has developed these resistors and they are appearing more frequently in modern electronic equipment. Color TV in particular has brought these resistors into considerably more importance.

If you have to replace one of these resistors, you should try to obtain an exact duplicate replacement. Not only is the cold resistance of a thermistor or varistor important, but also the way that the resistance changes either with changes in temperature in the case of the thermistor or changes in voltages in the case of the varistor is important. The manner in which the resistance changes is usually of even more importance than the cold value of the part.

High-voltage resistors are used in circuits where the length of standard resistors is so short that if you tried to use one the voltage would simply arc across the resistor.

These resistors are usually made by the spiralled-carbon method so that a long path can be obtained in order to get a very high resistance.

SELF-TEST QUESTIONS

(aa) What is a thermistor?
(ab) Where are thermistors used?
(ac) Draw the schematic symbol for a thermistor.
(ad) What is a varistor?
(ae) Where are varistors used?
(af) How do you tell the difference between a thermistor and a varistor?
(ag) What is a high-voltage resistor?

Meters

We have already mentioned that resistors made with tolerances of 1% and 1/2% are used in electronic test equipment. Resistors made to these close tolerances are often called precision resistors. Precision resistors are used with meters in order to increase the usefulness of a given meter movement. Since this represents another important use of resistors, and since meters are very important, we are now going to present some of the basic fundamentals about meters.

The meter most frequently used in electronics for current, voltage and resistance measurements is the d'Arsonval* meter. The meter gets its name from the scientist who invented it.

Basically the d'Arsonval meter is a current-indicating device, but by arranging it in suitable circuits it can be used to measure voltage and resistance as well as current.

Let us go ahead and learn how the meter itself works and then we will see what changes we can make in the basic meter to make it more usable.

*Pronounced dar-son-val. The dar is like ar in car, son like son or sun, and the val like all in ball. d'Arsonval.

THE BASIC METER MOVEMENT

A simplified drawing of a d'Arsonval type meter is shown in Fig. 16. Notice that we have a permanent magnet and that the faces of the poles of the magnet have been curved. Between the poles of the magnet we have a coil that is wound on a very light frame. Usually the frame is made of a thin piece of aluminum.

The frame and coil are attached to light-weight pivots, and these pivots fit into jewel bearings. The bearings are securely supported by the meter frame so that the coil and the pivots are held in position and can rotate freely between the poles of the magnet. Attached to each pivot and anchored to the meter frame is a spring. There is a spring on each pivot and these springs are arranged to rotate the coil into approximately the position shown in the figure. Also attached to the pivot is a pointer as can be seen in the drawing. When

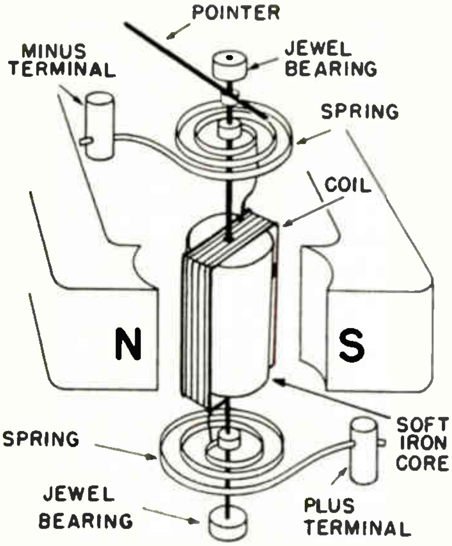


Fig. 16. How a d'Arsonval type of meter is made.

the coil rotates this pointer will rotate in a clockwise direction.

Notice that the ends of the coil are connected to the springs and the other ends of the springs are connected to terminals. One terminal is the negative terminal and the other terminal is the positive terminal.

When a current flows through the meter coil, the coil will become an electromagnet. If the current enters the negative terminal of the meter and leaves the positive terminal a field will be set up in the electromagnet that will be attracted by the field of the permanent magnet. The attraction between the two fields will produce a force called torque that will try to rotate the coil. The torque will cause the coil to rotate in a clockwise direction against the force or torque of the springs until the torque in the coil is balanced by the opposing torque of the springs. The higher the current flowing through the coil, the more torque is produced and the more the coil will be able to rotate against the opposing torque of the springs. As the coil rotates, the meter pointer moves in a clockwise direction and moves up the scale of the meter, indicating the current that is flowing through the coil.

A typical current scale is shown on the meter in Fig. 17. Notice that when the meter pointer moves all the way to the right side of the scale, the current flowing is 1 amp. We call this meter an ammeter because it measures the current in amps. We call it a 1 amp ammeter because a current of 1 amp gives a full scale deflection. If a current of 1/2 amp flows through the meter coil, only half as much torque, or rotating force will be developed, and as a result, the meter pointer will move only halfway up the scale. If the current flowing through the meter coil is one quarter of an amp, then one-fourth the torque needed to produce a full-scale deflection will be produced and the meter pointer will move only one-quarter of the way up the scale.

We mentioned earlier that the coil

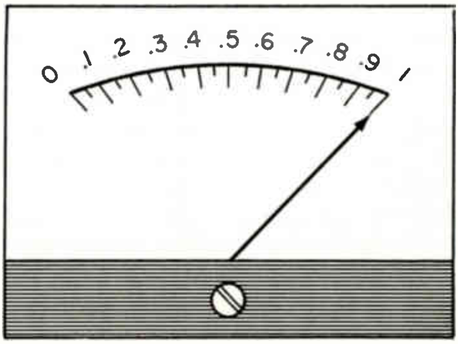


Fig. 17. A 1-amp ammeter.

was generally wound on a lightweight aluminum frame. The aluminum frame acts like a coil made of one turn of wire that is shorted. When a current flows through the meter coil and the coil begins to rotate, the one turn coil made up of the aluminum frame also rotates. As the aluminum frame rotates, it will cut through the field produced by the permanent magnet and a voltage will be induced in it. Although the voltage induced in the one turn coil made up of the frame will be low, a fairly high current will flow through the coil because it has a very low resistance. This in turn sets up a significantly strong magnetic field that opposes the motion of the meter coil. The net effect is that this opposition damps the movement of the coil. It prevents the coil from swinging too rapidly in a clockwise direction - this would cause the coil to swing past the position it is supposed to reach and oscillate back and forth around the position it should reach. The field produced by the current flowing in the aluminum frame will prevent this from happening and as a result the pointer will move upscale at a reasonable speed and indicate the value of current flowing with a minimum of oscillation back and forth past the correct value due to the coil rotating too rapidly when the current is first applied to it.

The d'Arsonval meter can be made very sensitive. By using a very lightweight frame and many turns of very fine wire to wind the coil and keeping the weight of the moving part of the meter as low as possible, it is possible to build a meter of this type that will indicate a current of 50 microamps quite easily. As a matter of fact, d'Arsonval meters that can indicate currents of only a few microamps can be made, but usually the most sensitive d'Arsonval meters used by technicians for routine measurements are 50 microampere meters. More sensitive meters are quite expensive and very delicate and are usually used in laboratory-type measurements rather than in the type of measurements that the service technician will normally make.

MEASURING CURRENT

To measure the current flowing in a circuit, the meter is placed in the circuit so that the current flowing in the circuit must flow through it. Fig. 18 shows how a meter can be connected in a simple series circuit to find the current flowing in the circuit. Notice the schematic symbol for the meter at B. The letter I, for current, may be placed inside of the circle as in the figure. Sometimes the letter A, meaning amperes, is placed in the circle and sometimes MA, meaning milliamperes, is used. μA is also used to indicate microamps.

Before you measure the current in a circuit, you need to have some idea of what the current might be. D'Arsonval meters are made in many different ranges. We already mentioned that this type of meter can be made very sensitive so that it can measure a current of only a few microamperes. If the meter is designed to measure currents in microamperes, it is called a microammeter. Some d'Arsonval meters have a scale on them that indicates the current in milliamperes. This type of meter is called a milliammeter. Still other meters are made to measure higher currents and the scale indicates the current in amperes; this type of meter is called an ammeter. All work on the same principle; they get their different names from the current values they are designed to measure.

If you should connect a microammeter in a circuit where the current flow is several amperes, the current flow through the meter would be so high that either you would burn out the meter coil or else the very strong current flowing through the coil would cause the coil to rotate so violently, that it would either bend the meter pointer by starting too fast or else slam the meter pointer up against the end of the scale and ruin the meter. Therefore, as we mentioned earlier you need to have some idea of what the current flow in the circuit is so you can be sure to use a meter that is capable of measuring the current flowing in the circuit. If you don't know what the current is, you should use the highest range meter you have. You can always put a more sensitive meter in the circuit later, if you find that the current is small and will not damage a more sensitive meter.

We mentioned that meters are available in many different ranges. Actually, meter manufacturers do not make a large number of different meter movements. Instead, they make a few standard meter movements and then extend the range of the meter by means of shunts.

You have already studied parallel circuits and you know that when two resistances are placed in parallel, part of the current will flow through one resistance and part through the other. The coil in a d'Arsonval meter has a certain resistance. For example, if you have a meter that indicates a current of 10 milliamperes and the coil has a resistance of 100 ohms, and if you place a 100-ohm resistor in parallel with the meter, half the current flowing in the circuit would flow through the meter

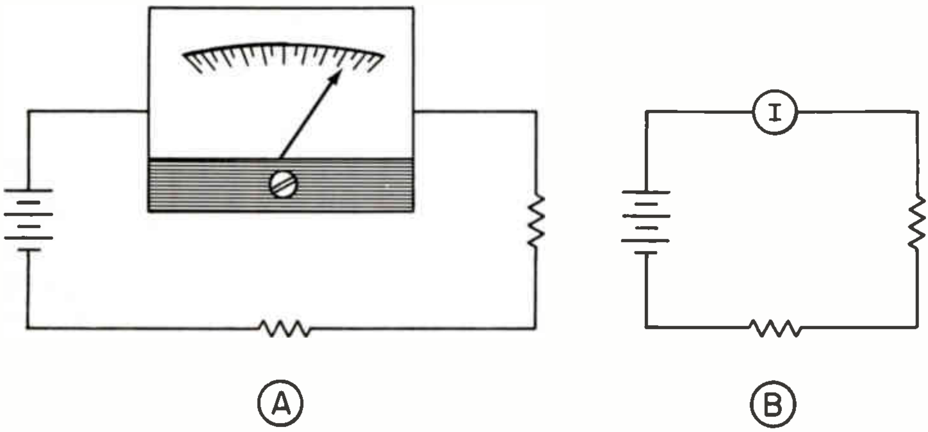


Fig. 18. A meter connected into a series circuit to measure current. Pictorial of meter is shown at A; schematic symbol, at B.

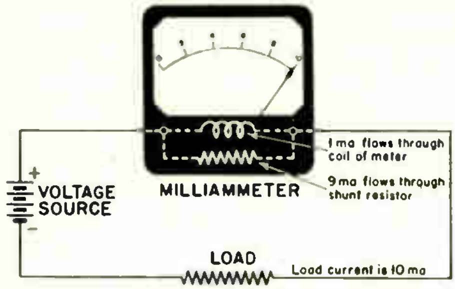


Fig. 19. By adding a shunt resistor, this 0-1 ma dc milliammeter can be made to read current up to 10 milliamperes.

coil and half through the resistance. In this case, if the total current is 10 milliamperes, only 5 milliamperes will flow through the meter coil. The remainder will flow through the shunt resistor. This means that the meter pointer would move only to half-scale.

An example of a meter used with a shunt is shown in Fig. 19. Here the basic meter movement is designed to give a full-scale deflection with a current of 1 milliampere through the coil. The meter has a resistance of 100 ohms. A shunt has been placed across the meter coil having one-ninth the resistance of the coil. Now, the current divides so that nine-tenths of the current will flow through the shunt and one-tenth through the coil. Therefore, it is possible to place a scale on the meter that indicates 10 milliamperes at full scale. When the total current flowing is 10 milliamperes, 1 will flow through the coil and cause the meter to read full scale and 9 will flow through the shunt.

Manufacturers use shunts to provide a large number of meters with different ranges. The chances are if you buy a 50 ma meter and a 500 ma meter, they will both be either 5 ma or 1 ma meters with suitable shunts inside the meter to give the higher ranges.

DC VOLTMETERS

A d'Arsonval meter can be used to read voltages by connecting a resistor in series with the meter. This resistor is called a multiplier. In Fig. 20 we have shown how a d'Arsonval meter can be used with a multiplier to measure voltage. The basic meter is a 1 milliampere meter. This means that at full scale the meter will indicate a current of 1 milliampere. Now, you know from Ohm's Law that when a voltage is applied to a resistance a current will flow through the resistance and the value of the current will depend upon the voltage and on the resistance. If we put enough resistance in series with the meter so that the total resistance is 1000 ohms, then it will take a voltage of 1 volt to cause a current of 1 ma to flow through the resistance and the meter. If we connected the combination across a bat-

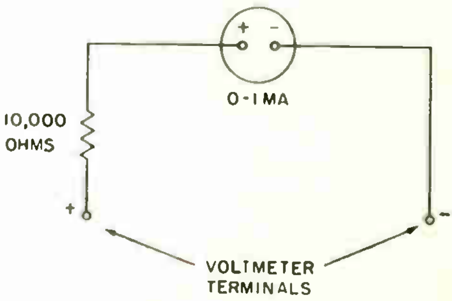


Fig. 20. A dc voltmeter made from a 0-1 ma dc milliammeter and a 10,000-ohm resistor.

tery and the battery voltage was 1 volt, a 1 milliampere current would flow and the meter pointer would indicate full scale. If the voltage was only 1/2 volt, then only half a milliampere would flow in the circuit and the meter pointer would move up to half scale.

If instead of a 1000-ohm resistance we add enough resistance in series with the meter to give us a total resistance of 10,000 ohms, then it would take 10 volts to cause a current of 1 milliampere to flow in the circuit. Under these conditions the meter pointer would again move up to full scale.

With a multiplier in series with the meter we can calibrate the meter directly in volts rather than in milliamperes. With a total resistance of 10,000 ohms in the circuit we can mark the full scale reading as 10 volts and half scale reading as 5 volts and so on. If the total resistance in the circuit is 100,000 ohms then a full scale reading would be 100 volts.

Notice that for each volt we have added a resistance of 1000 ohms in series with the meter. If we want the meter to read 1 volt full scale then we need a total resistance of 1000 ohms. If we want the meter to read 10 volts full scale then we need a total resistance of 10,000 ohms and if we want the meter to read 100 volts full scale then we need a total resistance of 100,000 ohms. We call this type of meter a 1000-ohms per volt meter. We say that the meter sensitivity is 1000 ohms per volt. You will see later that the sensitivity of the meter is important in taking voltage measurements.

The meter we have been discussing so far has been a 1 milliampere meter. However, if instead of using a 1 milliampere meter as a voltmeter we start with a basic meter movement of 50 microamperes we will have a much more sensitive meter. If we convert a 50 u amp meter to a voltmeter when the meter pointer reads full scale, it means that the current flowing through the meter and its multiplier is 50 microamperes. If we want to build a meter that will indicate full scale when a voltage of 1 volt is applied to it, we can use Ohm's Law to find out how much resistance we need in the circuit. Using

$$R = \frac{E}{I}$$

and substituting 1 volt for E and 50 microamperes for I we get

$$R = \frac{1}{.000050}$$

and we can multiply by 1,000,000 to convert this current to amperes and then we will have

$$R = \frac{1 \times 1{,}000{,}000}{50}$$

$$R = \frac{1{,}000{,}000}{50}$$

and now cancelling one zero above and below the line, and dividing 5 into 100,000 we find that

$$R = 20{,}000 \text{ ohms}$$

We say that the sensitivity of this meter is 20,000 ohms per volt. If instead of a 1-volt meter we want a 10-volt meter we need to add ten times as much resistance or a total of 200,000 ohms. If we want a 100-volt meter then we would have to add 100 × 20,000 or 2,000,000 ohms in series with the meter.

Since the higher sensitivity meter actually draws less current from the circuit, the meter will have less tendency to upset the circuit when voltage measurements are taken. This is particularly important in electronic circuits where the resistance is high. If you use a 1,000-ohm per volt meter to take a voltage measurement, the chances are that the meter itself will take more current than is actually being used to operate the circuit. This means that the meter will upset the circuit so that the voltage reading you will obtain will not be the actual voltage that is present in the circuit when the meter is disconnected. On the other hand, a 20,000-ohm per volt meter has twenty times the resistance; it takes far less current from the circuit and is much less likely to upset the voltages in the circuit.

OHMMETERS

An ohmmeter is a device used to measure resistance. Actually, an ohmmeter is nothing more than a d'Arsonval type meter used in conjunction with a battery and a series resistor. A circuit of a typical ohmmeter is shown in Fig. 21.

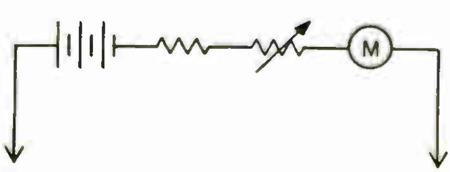


Fig. 21. A schematic diagram of an ohmmeter.

As you can see in the circuit the ohmmeter consists of a 4.5 volt battery, two resistors (one a fixed resistor and the other a variable resistor), and a 1 ma meter. The fixed resistor has a resistance of 4000 ohms. The adjustable resistor has a resistance of 1000 ohms and under normal circumstances you would short the test probes of the ohmmeter together and then adjust the variable resistor to get a full scale reading. If the battery voltage is exactly 4.5 volts, you will get a full scale reading, in other words, a current of 1 ma through the meter when the resistance of the potentiometer is set at slightly less than 500 ohms. With this setting of the potentiometer you will have a total resistance in the circuit of 4500 ohms. The potentiometer is set at a value slightly less than 500 ohms to make up for the internal resistance of the battery and the resistance of the meter. In any case, with a voltage of exactly 4.5 volts, the resistance in the circuit will be 4500 ohms.

Now if you separate the test probes, the circuit will be open and the meter pointer will drop back to 0. If you place a 4500-ohm resistor between the test probes, then the total resistance in the circuit will be 9000 ohms, and the current that will flow in the circuit will be .5 ma. In other words, you will get a half-scale deflection on the meter. If you were measuring the resistance of an unknown value and you got a half-scale reading, you would know immediately that its resistance was 4500 ohms.

As a matter of fact, you can take any current reading that you might get on the meter scale and use Ohm's

Law to calculate the total resistance in the circuit. Knowing that the battery voltage is 4.5 volts, you simply use the formula

$$R = \frac{E}{I}$$

and substitute 4.5 volts for the battery voltage and the value of current indicated on the meter. This will give you the total resistance in the circuit. From this value you subtract 4500 ohms, which is the fixed resistance in the circuit and this will give you the resistance of the unknown resistor.

Instead of going through this procedure of calculating the resistance value each time you make a measurement, these calculations can be worked for various current values and the meter calibrated directly in ohms. This is what is done in an ohmmeter.

If you need to measure higher resistances, you can use a higher battery voltage. For example, if you used a 45-volt battery then you have to put a total of 45,000 ohms in the circuit to get a full-scale meter reading. Then a center scale reading would represent a resistance of 45,000 ohms, instead of 4500 ohms as before.

Another way of building a meter that will give higher resistance readings conveniently is to use a more sensitive meter. If you use a 50 microamp meter, you will need a total resistance in the circuit of 90,000 ohms to give you a full-scale reading. Thus with this type of meter, a center scale reading will indicate a resistance of 90,000 ohms.

MULTIMETERS

A multimeter is simply a meter in which a single d'Arsonval type meter is arranged with a series of switches, resistors and batteries, so that by rotating the switches it can be used to perform a large number of functions. One switch is usually called the function switch. By putting this switch in the correct position, you can use the meter either to measure voltage, current or resistance. The other important switch on the multimeter is called the range switch. This switch controls the full-scale reading that you will obtain on the meter. In other words, with the switch in one position, when you are measuring voltage, a full-scale reading might indicate 10 volts. With the switch in another position the full-scale reading might indicate a voltage of 100 volts, and in a third position it might indicate a voltage of 500 volts.

Multimeters are widely used by service technicians because the basic d'Arsonval meter movement is quite expensive, and it is more economical to use a single meter with suitable switches and resistors to perform all types of measurements than it is to have a number of separate meters for different measurements.

SUMMARY

As an electronics technician, you will not have to know how meters are designed or built. You should understand the basic operation of a meter so that you will know how it works - this will help you use it to the best advantage. You must know that meters are very delicate and must be handled carefully. If you drop a meter or bang it hard you are liable to knock the pivots out of the jewel bearings, in which case the meter pointer will stick as it moves upscale.

You must avoid overloading a meter. The coiled springs in a meter are made of phosphor bronze. If they are overloaded they overheat and their shape is distorted. If adjacent turns of a spring touch enough, friction is produced to upset the meter accuracy.

Technicians seldom have to take current measurements, but you should know how to connect a meter in order to take a current measurement. You must remember to use a meter that is capable of measuring the current in a circuit. If you have no idea what the current is you should start with the largest meter you have first and then work down to a smaller range meter after you are sure that the current is low enough not to damage it.

Voltage measurements are probably the most important measurement to the technician. You need to know how to take a voltage measurement and you need to understand what effect the meter sensitivity will have on the accuracy of the voltage measurement. A meter with high sensitivity will take less current from the circuit and there will be less chance of the meter upsetting the circuit so that you will obtain an erroneous voltage measurement.

When you start working on your kits, you will use your meter to take both current and voltage measurements. You will learn how to read the scale on the meter - we will not go into it in this lesson because it will be much easier to learn to read the scale when you actually have the meter in front of you.

SELF-TEST QUESTIONS

(ah) What is the name given to the basic meter movement most frequently used in voltage, current and resistance measurements?

(ai) What causes the pointer of a meter to move upscale when a current flows through the meter?

(aj) What force must the coil in a meter overcome before it can rotate?

(ak) What purpose, other than to hold the coil, does the aluminum frame on which the coil is wound serve?

(al) How is a meter connected to measure current flowing in a circuit?

(am) How do you connect a voltmeter in order to measure the voltage across a part?

(an) Which type of voltmeter will upset the circuit performance less, a 1000-ohm-per-volt meter or a 10,000-ohm-per volt meter?

ANSWERS TO SELF-TEST QUESTIONS

(a) A series-dropping resistor drops the available voltage to the value required by the load.

(b) 38 volts. The two 35-volt tubes will require a total voltage of 70 volts in series, and the two 6-volt tubes will require an additional 12 volts, giving a total required voltage of 82 volts. Subtracting 82 volts from the 120-volt power line gives us a voltage of 38 volts which the series-dropping resistor must drop.

(c) 126.6 ohms. Using Ohm's Law to find the value resistor we have a voltage of 38 volts and a current of .3 amps. Using the formula

$$R = \frac{E}{I}$$

and substituting the values for current and voltage we get

$$R = \frac{38}{.3} = 126.6 \text{ ohms}$$

Since it would be impossible to obtain a resistor of this value you use the nearest standard size which is 130 ohms. If you get 126 or 127 ohms as your answer your answer is close enough; when you went to buy a resistor you would simply have to take a resistor having a resistance as close as you could get to the calculated value.

(d) A bleeder is used to regulate or help maintain constant the voltage across a load.

(e) A bleeder that draws a substantial current will be more effective than a high-resistance bleeder that draws only a small current. The greater the bleeder current in proportion to the load current, the more effective the bleeder will be insofar as regulating the load voltage is concerned.

(f) The current consumed by the bleeder is wasted current insofar as performing any other useful function other than regulating the voltage across the load is concerned. Since the more current you draw from a power supply, the larger the components (hence the more expensive they must be), the bleeder current should be no larger than necessary to give you the regulation required. Excessive bleeder current with regulation better than is required is costly and can run the cost of the power supply substantially higher than necessary.

(g) 10K ohms. You can determine the resistance of the series-dropping resistor two ways. The easy way: if you notice that the voltage across the load is 200 volts and you have to drop 100 volts across the dropping resistor, you can see that the resistance of the dropping resistor must be half the resistance of the load. The other way is to calculate the current through the load resistor using Ohm's Law. Once you have the current you can then calculate the size resistor required as the series-dropping resistor to drop 100 volts. From Ohm's Law the current through the load will be

$$\frac{200}{20{,}000} = .01 \text{ amps}$$

Now to find the value of the dropping resistor you have to use the formula

$$R = \frac{E}{I}$$

$$= \frac{100}{.01} = 10K \text{ ohms.}$$

(h) The watt.

(i) 45 watts. You use the formula

$$P = E \times I$$

and substituting 15 for E and 3 for I we get

$$P = 15 \times 3 = 45 \text{ watts.}$$

(j) When we say a resistor is dissipating 10 watts, we mean that the resistor is using 10 watts of electrical energy. We say it is dissipating the power because it changes the power from electrical energy to heat. It is taking 10 watts of electrical energy out of the circuit and converting it to heat.

(k) 10 watts. To find the power dissipated by the resistor we use the formula

$$P = \frac{E^2}{R}$$

Remember that $E^2 = E \times E$. Substituting 100 for E and 1000 for R we have

$$P = \frac{100 \times 100}{1000}$$

and cancelling three zeros above the line and three zeros below the line we have:

$$P = 10 \text{ watts.}$$

(l) 1/2 watt. Again, to solve this problem we use the formula:

$$P = \frac{E^2}{R}$$

Substituting 50 volts for E and 5000 ohms for R we have

$$P = \frac{50 \times 50}{5000} = \frac{2500}{5000} = 1/2 \text{ watt.}$$

(m) 100 watts. To solve this problem we use the formula

$$P = I^2R$$

substituting 2 for I and 25 for R we have

$$P = 2 \times 2 \times 25 = 100 \text{ watts.}$$

(n) 10 watts. To solve this problem we must convert milliamps to amps and K-ohms to ohms and then use the formula

$$P = I^2R$$

10 ma = .01 amps.
100K = 100,000 ohms.

Substituting these values in the formula we get

$$P = .01 \times .01 \times 100{,}000$$
$$P = .0001 \times 100{,}000.$$

To multiply 100,000 by .0001 we simply move the decimal point four places to the left thus

$$P = 100{,}000 \times .0001 = 10 \text{ watts.}$$

(o) 3 ohms. Maximum power transfer will be obtained when the load resistance is equal to the battery resistance.

(p) You would use a 10-ohm resistor to get maximum power transfer to the load; resistance should equal the battery resistance. With a 10-ohm load resistance, the total resistance in the circuit will be 20 ohms, and therefore the current flow in the circuit will be

$$I = \frac{20}{20} = 1 \text{ amp.}$$

The total power produced by the battery will be

$$P = E \times I$$
$$P = 20 \times 1 = 20 \text{ watts.}$$

The power transferred to the resistor can be found using the formula

$$P = I^2R$$
$$P = 1 \times 10 = 10 \text{ watts.}$$

(q) 4700 ohms.

(r) .39 megohms equals 390K-ohms; .39 megohms is also equal to 390,000 ohms.

(s) 680,000 ohms equals 680K equals .68 megs.

(t) The resistor is within its rated tolerance. A 2200-ohm resistor with a tolerance of 10% may vary as much as 220 ohms above or below its indicated value. Subtracting 220 from 2200 gives us 1980 as the lower limit of the resistor. Since 2000 ohms is between this lower limit and the indicated value of the resistor, the resistor is within tolerance.

(u) 10,500 ohms. 5% of 10,000 ohms is 500 ohms. Therefore the maximum value that the resistor can have and still be within tolerance will be 10,000 plus 500 equals 10,500 ohms.

(v) 390,000 ohms, which is also equal to 390K-ohms. The tolerance is 5%.

(w) 1,000,000 ohms or 1 megohm. The tolerance is 10%.

(x) 56,000 ohms or 56K-ohms. The tolerance is 5%.

(y) 2200 ohms or 2.2K-ohms. Its tolerance is 10%.

(z) Yes. The 1/2% resistor has a closer tolerance than the 1% resistor. This means that it will vary less from the indicated value than the 1% resistor - in other words it is a better resistor. You can always use a better resistor - one with a closer tolerance as a replacement.

(aa) A thermistor is a special resistor with a negative temperature coefficient. In other words, as the resistor heats, its resistance will decrease.

(ab) Thermistors are used in circuits in which high current surges are often obtained when the equipment is first turned on. The high cold resistance of the thermistor limits the initial current surge. As the thermistor heats up, the resistance drops so that it has very little effect on the circuit performance.

(ac) See Fig. 12.

(ad) A varistor is a voltage-dependent resistor. A voltage-dependent resistor is a resistor whose resistance depends upon the voltage applied across it. If the voltage increases, the resistance of the varistor decreases.

(ae) Varistors are used in circuits where sudden increases of voltage could damage components or otherwise upset the performance of the circuit. With a varistor in the circuit, as the voltage increases, the resistance of the varistor decreases, drawing a large current from the voltage source. The increased current drain tends to lower the voltage to a safe level.

(af) Often you cannot tell the difference between a thermistor and a varistor simply by looking at them. You should refer to the schematic diagram of the equipment in which they are used to identify each type.

(ag) A high-voltage resistor is a resistor made for use in high voltage circuits. The resistor is longer than most resistors in order to prevent a voltage arc across the resistor. High-voltage resistors are usually made by the spiral-wound carbon technique in order to provide a carbon path long enough to produce the required resistance.

(ah) The d'Arsonval meter.

(ai) The current flowing through the meter coil sets up a magnetic field. This field working with the magnetic field of the permanent magnet in the meter produces a torque which causes the coil and its frame to rotate. The meter pointer is attached to the coil frame and so the meter pointer moves upscale.

(aj) The opposing torque of the springs. The springs are used to hold the meter coil and pointer in a 0 position when no current flows through the coil, and to oppose the movement of the coil when current flows through the coil.

(ak) The aluminum frame acts as a shorted turn. The motion of the coil in the magnetic field of the permanent magnet induces a voltage in the coil which opposes the movement of the coil and frame. This tends to damp the coil movement and prevent its oscillating or swinging back and forth past the indicated value.

(al) The meter is connected in series with the circuit to measure current flowing in the circuit.

(am) You connect the voltmeter directly across the part where you want to measure the voltage. The voltmeter must be connected with the proper polarity; the negative side of the meter must be connected to the end of the part that is connected closest to the negative terminal of the voltage source, and the positive terminal must be connected to the end of the part closest to the positive terminal of the voltage source.

(an) A 10,000-ohm-per-volt meter will have less effect on the circuit because it will take less current from the circuit when you take a voltage measurement.

Lesson Questions

Be sure to number your Answer Sheet B105-1.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable. However, don't hold your answers too long; you may lose them. Don't hold answers to more than two sets at a time or you may run out of lessons before new ones can reach you.

1. Name three important kinds of wire-wound resistors.

2. What is a series-voltage-dropping resistor?

3. A 750-ohm load must be operated with a voltage of 75 volts. The output voltage from the power supply is 125 volts. What value of series-dropping resistor should be used to drop the power supply voltage to the correct value for the load?

4. A 15-volt battery is supplying a current of 3 amps to a load. What is the power the battery is supplying?

5. The voltage across a 5000-ohm resistor is 100 volts. How much power is the resistor dissipating?

6. What is the purpose of a bleeder?

7. What is a thermistor?

8. Express the following resistance values in ohms: (a) 2.2K, (b) 470K, (c) 3.3 megs, (d) .17 megs. (e) .47 megs.

9. An ammeter has a full scale reading of 1 amp and an internal resistance of 1 ohm. A separate external 1-ohm shunt is connected in parallel with the meter. If the meter and shunt are connected into a circuit, and the meter indicates a current of .5 amps, what current is actually flowing in the circuit?

10. What do we mean when we say that a meter has a sensitivity of 1000 ohms per volt?

A PLAN FOR YOUR FUTURE

In a radio interview a few minutes after a championship heavyweight boxing match, one of the fighters stated his plans for the future as follows:

> "I'm going to get myself in shape, fight my own fights, and listen to nobody!"

You can use these dynamite-packed words as your plan for the future, too. Here's the way:

"GET MYSELF IN SHAPE." You're doing this right now, because the NRI Course gets you in shape for a career in electronics. But remember that it takes the complete NRI Course, with all its associated practical work, to get you completely in shape.

"FIGHT MY OWN FIGHTS." In real life, the only person who can bring you success is YOU yourself. Expecting somebody else to do your work and fight for your success is just wishful thinking.

"LISTEN TO NOBODY." Even friends and relatives will at times ridicule your studies -- they can't help it, because seeing you get ahead makes them feel uncomfortable about their own laziness. So, remember human nature, and don't give anyone a chance to discourage you.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-205

ACHIEVEMENT THROUGH ELECTRONICS

nri

HOW COILS ARE USED

B106

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

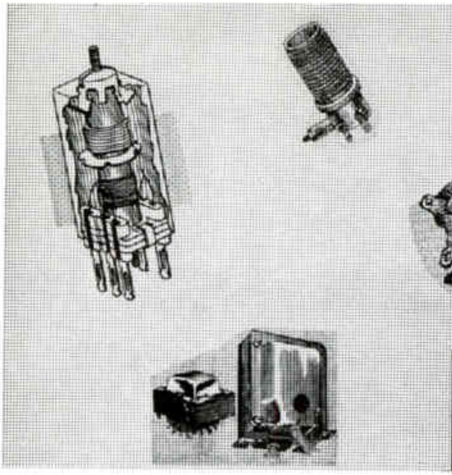
HOW COILS ARE USED

B106

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

JDL3M1174 1974 EDITION Litho in U.S.A.



HOW COILS ARE USED

Coils are important to the electronics technician because they are used in all types of electronic equipment. They are used in many different ways to perform different jobs. Many different types of coils are used in electronic equipment - some coils may have only one or two turns whereas other coils may have several thousand turns. You will find all types of coils in industrial electronic equipment, in radio and television receivers, and for that matter, in practically every type of electronic equipment you will encounter. Since you will find coils used in all types of electronic equipment, it is important for you to understand what they are used for and how they work.

HOW COILS ARE USED

During the daytime you can probably find somewhere between ten and twenty-five different stations operating on the standard broadcast band on your radio receiver. During the evening hours you can probably find even more stations coming in. Coils are used with capacitors in special circuits called resonant circuits to enable you to select one signal and reject the signals from the other radio stations. Similarly, coils are used in television receivers to enable you to tune from one channel to another. The coils used for this purpose in radio receivers will have many turns on them whereas the coils used to select the different channels in a TV receiver will have only a few turns on them.

Coils are used in power supplies to help smooth the pulsating dc output from a rectifier to pure dc. Coils used for this purpose are called filter chokes or simply chokes. They act to permit direct current to flow through them without any opposition, but offer a high opposition to ac current or any change in the amplitude of the direct current.

Coils are used to produce motion. You have already seen an example of this in the d'Arsonval meter. The current flowing through the coil produces a magnetic field which acts with the field from the permanent magnet to cause the coil to rotate. Coils are used in motors in much the same way; the magnetic field produced in a coil wound on a form called an armature opposes or attracts the field produced by a stationary magnet and the armature rotates. Coils are also used in generators to produce electricity; rotating the coil so that the turns of the coil cut through the magnetic lines of force produced by the field or stationary magnet of the generator results in a voltage being induced in the turns of the coil.

Coils are used in transformers. You already know that a transformer is nothing more than two or more coils wound on a common core. If ac is fed to one of the coils, which we call the primary coil, a voltage will be induced in the other coil, which we call the secondary coil. If the secondary coil has more turns on it than the primary coil, the total voltage produced by the secondary coil will be higher than the primary voltage, but if the secondary coil has fewer turns on it than the primary coil, the voltage produced by the secondary winding will be lower than the voltage applied to the primary coil.

TYPES OF COILS

We have already mentioned that a coil used to select the various channels in a TV receiver may have only a few turns. In fact, a coil used in a UHF TV tuner is often made with a wire shaped something like a ribbon rather than a round piece of wire. The material is usually silver plated to cut down losses and the total coil will consist of about three quarters of a turn with a diameter of about an inch and a half. On the other hand, in a long wave receiver (a receiver designed to receive stations lower in frequency than the standard radio broadcast band) you may find coils wound on a form about an inch in diameter and having close to one thousand turns. Both coils perform the same task - they are used along with capacitors to select one radio frequency signal and reject others.

We mentioned that choke coils were used in power supplies to help provide smooth dc. A choke coil of this type found in a typical color TV receiver will be wound on an iron core. The coil itself will usually have several hundred turns and the complete assembly may weigh two or three pounds. On the other hand, the choke coil used for the same purpose in a large radio transmitter will be much larger. The iron core itself will be much larger and a much larger size of wire will be used because the wire will have to carry a heavier current. Chokes of this type will weigh one hundred pounds or more.

You will also run into small chokes called radio frequency chokes (abbreviated rfc). A radio frequency choke may have only a few turns and may be completely self-supporting or it may have several hundred turns and be wound on some nonmagnetic type of material. The number of turns required on a radio frequency choke depends on the frequency at which the choke is to be used and how much opposition it must offer to the flow of radio frequency currents through it. Generally speaking, the higher the frequency at which the choke is to be used, the fewer the number of turns the choke will have.

In spite of the great difference in the size of the coils we have described, their operation is basically the same. Thus it is important for you to understand the basic facts about how a coil works. Once you have learned these facts you will know exactly what a coil does in the circuit whether it is a small coil having two or three turns or a large coil with many turns wound on an iron core.

COIL ACTION

In its simplest form a coil is nothing more than one or two turns or loops of wire, usually wound in a circular or helical (spiral) shape. If the coil is self-supporting so that it has no core other than air, it is called an air-core coil. If the coil is wound on a cardboard, ceramic, or non-magnetic type of material, it is also called an air-core coil. The cardboard form or ceramic form are used only to hold the turns of wire in place, and have no appreciable effect on the performance of the coil. On the other hand, if the coil is wound on a form made up of a magnetic type of material such as iron, it is called an iron-core coil. Air-core coils are frequently placed inside metal shields or housings to prevent their picking up interference from outside sources or producing interference in other circuits. These shields may have some effect on the performance of the coils, but it is quite different from the effect of an iron-core and so they are still considered as air-core coils.

You already know that when a current is flowing through a wire there is a magnetic field around the wire as shown in Fig. 1A. The current produces magnetic lines of force, which are also known as magnetic flux or flux lines. When the wire is bent into a loop as shown in Fig. 1B, the magnetic lines of force all go through the loop in the same direction. This concentrates the magnetic flux around the loop. If, instead of being bent into a single loop, the wire is bent into a number of loops so that the coil has a number of turns, additional magnetic flux will be created, which will result in a stronger magnetic field.

Because of the magnetic flux, the coil offers more opposition to the flow of alternating current than it does to the flow of direct current. You will see in this lesson why this is so. Once you understand why a coil offers more opposition to ac than to dc, you will have a good understanding of how a coil works.

Since the operation of a coil depends upon the magnetic flux produced by the coil, it is important that you know more about magnetic circuits. Therefore, before going ahead with our study of coils, we will study magnetic circuits.

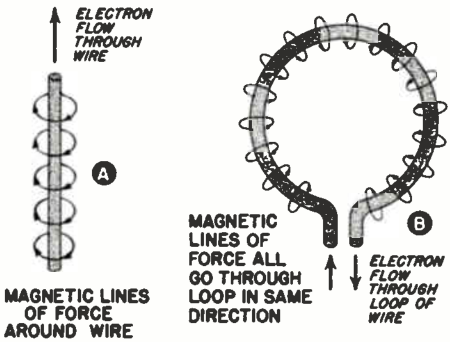


Fig. 1. When current flows through a wire, magnetic lines of force are set up around the wire as at A. When the wire is bent into a loop, the lines of force all go through the loop in the same direction, as shown at B.

Magnetic Circuits

You already know that the magnetic lines of force produced by a current flowing through a coil exist only as long as current flows. Once the current flowing through the coil stops, the magnetic lines of force will disappear.

Another important fact is that the magnetic lines of force are complete loops, having no ends. Notice that in Fig. 1, the magnetic lines of force around the wire and around the single coil were represented as complete loops. When two turns of wire are placed close together, the magnetic lines of force form complete loops around the two turns as shown in Fig. 2. Fig. 3 shows a coil made up of a number of turns and shows how the lines of force come out of the left of the coil and circle around to

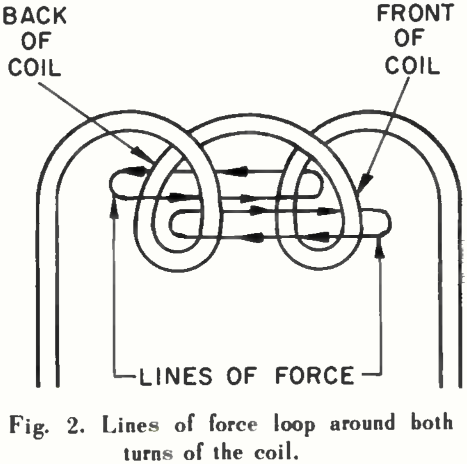


Fig. 2. Lines of force loop around both turns of the coil.

the right end of the coil. Notice that all of the lines of force are actually complete loops. We have not tried to draw all the magnetic lines of force that exist around the coil having a current flow through it. The

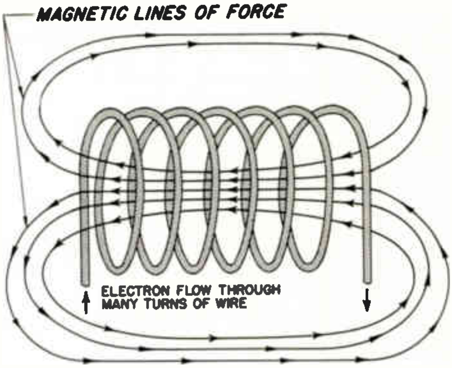


Fig. 3. When current flows through a coil, magnetic lines of force are produced. These are complete loops, passing through and around the coil.

coil may have thousands of such magnetic lines each forming a complete loop and passing through all or part of the coil and radiating out from the ends of the coil. The path of these magnetic lines of force is the magnetic circle of the coil. Although most of the lines of force will be concentrated near the coil, some will extend quite some distance from it.

AIR-CORE COILS

As we mentioned previously, some coils are completely self-supporting or wound on a cardboard or ceramic form. The purpose of these forms is simply to support the turns of wire making up the coil. Cardboard or ceramic serves no useful purpose as far as the operation of the coil is concerned. The actual core of the coil is simply air. This means that the lines of force must travel through air. This is the basic definition of an air-core coil: it is a coil in which the lines of force must travel through air or some other material that acts just like air insofar as the magnetic lines of force are concerned.

IRON-CORE COILS

Frequently a coil is wound on an iron or steel form or on a cardboard form with an iron or steel slug inside it. The iron or steel core makes a better magnetic circuit than air, and there will be a greater number and concentration of the magnetic lines of force. This means there will be more flux lines produced by the coil. This type of coil is called an iron-core coil. In coils designed for use in a power line or at audio frequencies, the iron-core is made up of thin strips of iron or steel called "laminations." These pieces of steel

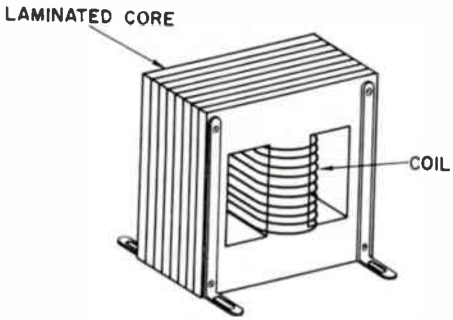


Fig. 4. Iron cores are frequently made of thin sheets of metal bolted together.

are fitted together to make a core, as shown in Fig. 4. As you can see, the core actually surrounds the coil and goes through the center as well. Fig. 5 shows how such a coil is constructed. This type of construction is used rather than a solid iron or steel core because it is more effi-

Fig. 5. Construction of an iron-core coil with laminated core.

cient. You will learn more about this in a later lesson.

In high-frequency circuits a powdered iron core is often used. This type of core is made by pulverizing iron filings and mixing them with a binder to hold them together and to insulate the particles from each other. A core of this type, called a slug, is often inserted in a coil by means of a screw so it can be adjusted in and out of the coil. This type of coil is shown in Fig. 6. The schematic symbol is shown beside it. Two or three lines drawn beside a coil indicate that it has an iron core, and the arrow indicates that it is movable. A coil with this type of core is called a slug-tuned coil.

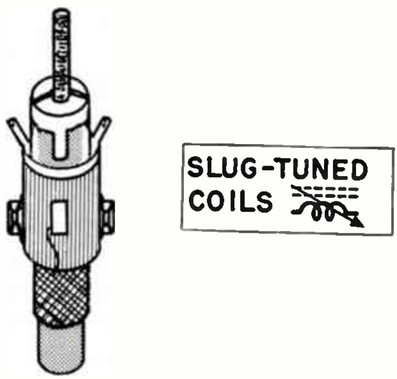


Fig. 6. A slug-tuned coil and its schematic symbol.

MAGNETOMOTIVE FORCE

The force that sends current around an electric circuit is called an electromotive force or voltage. The force that sends magnetic flux around a magnetic circuit is called magnetomotive force. It exists in every current-carrying coil.

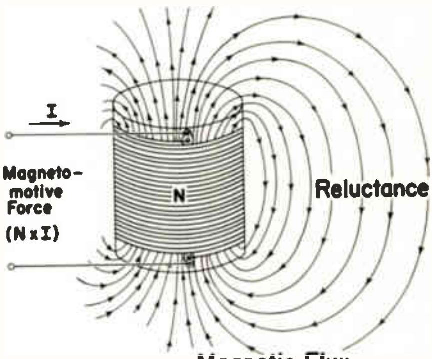
The unit of magnetomotive force is the ampere-turn. If a coil has one turn and the current flowing through it is 1 amp, the magnetomotive force is 1 ampere-turn. If the coil has 10 turns and the current flowing is 1 amp, the magnetomotive force is the product of the two, or 10 ampere-turns. If a coil has 5 turns, and the current flowing through it is 5 amperes, the magnetomotive force is 25 ampere-turns. Thus, to find the magnetomotive force in a coil, you simply multiply the current flowing through it in amperes by the number of turns on the coil. You can increase the magnetomotive force of the coil by increasing the current flowing through the coil or by adding more turns to the coil and keeping the current constant.

As an electronics technician it is very unlikely that you will ever have to calculate the magnetomotive force produced by a coil in a circuit. However, you should know what it is. It is essential to understand magnetic circuits in order to understand magnetic devices.

RELUCTANCE

You already know that every electric circuit has resistance. Resistance is the opposition to current flow in the circuit. Just as there is opposition to current flow in an electric circuit so also is there opposition to flux in a magnetic circuit. This opposition to flux is called reluctance.

The reluctance in a magnetic circuit is distributed along the entire path taken by the flux. In other words, there is reluctance all the way around the path followed by each magnetic line of force both in the core and in the air. We can actually make a very close comparison between a magnetic circuit and an electric circuit. In Fig. 7A we have the magnetomotive force sending the flux around the magnetic circuit against the opposition offered by the reluctance of the circuit. In Fig. 7B, we have a wire connected across a battery. Here the electromotive force of the battery is forcing current around the circuit against the opposition or resistance of the wire.



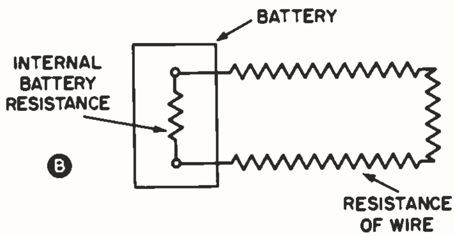


Fig. 7. Comparison between a magnetic circuit (A) and an electrical circuit (B).

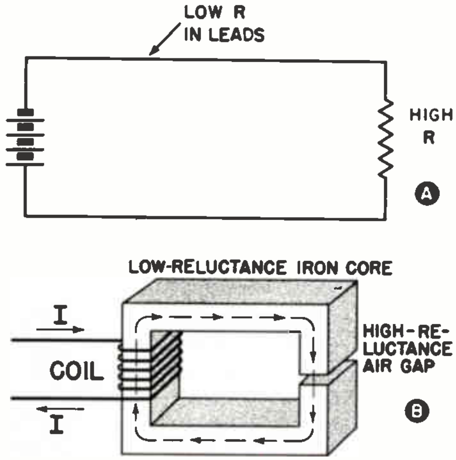


Fig. 8. In the electrical circuit at A, most of the opposition (resistance) to the current is in the resistor. The total resistance in the circuit is only slightly higher than the resistance of the resistor. In the magnetic circuit at B, most of the reluctance or opposition to the flux is in the air gap. The total reluctance in the circuit is only slightly higher than the reluctance of the air gap.

We have shown the whole length of the wire as a resistor. Here we see that the magnetomotive force is the equivalent of the electromotive force, the magnetic flux the equivalent of current, and the reluctance the equivalent of resistance.

We have a magnetic circuit that is similar to an electric circuit consisting of a resistor connected across a battery as shown in Fig. 8A. Here most of the resistance in the circuit is concentrated in the resistor; the resistance of the leads is small compared to that of the resistor. In the magnetic circuit of Fig. 8B, the iron core of the coil has an air gap. Most of the reluctance is concentrated in the air gap; the reluctance of the iron core is low in comparison to the reluctance of the air gap. One of the choke coils used in the power supply of a radio transmitter often has an air gap like this.

In an electric circuit, if we lower the resistance or opposition we can increase the current, and if we increase the resistance we reduce the current. Exactly the same situation exists in the magnetic circuit. If we lower the reluctance or opposition we increase the flux, and if we increase the reluctance we reduce the flux.

The reluctance in a magnetic circuit can be reduced by providing a better path through which the magnetic lines of force can flow. Materials such as iron and steel have a low reluctance, just as copper has a low resistance in an electric circuit. Therefore if a coil is wound on an iron core shaped like the one shown in Fig. 9, the magnetic lines will flow through the core as shown in the drawing. Because the iron has a low reluctance, there will be a much greater flux than there would

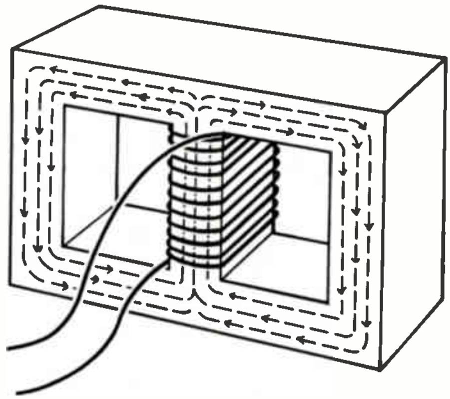


Fig. 9. There will be much greater flux if a coil is wound on an iron core like this than if the coil has an air core.

be if the same coil had an air core. Thus the flux in a magnetic circuit can be increased by providing a path of a magnetic material through which the magnetic lines of force can flow. Making a frame like the one shown in Fig. 9 of nonmagnetic material such as paper, glass, aluminum, or copper would not increase the flux, because these materials do not have a lower reluctance than air.

There are other factors that affect reluctance. Increasing the cross sectional area of the core will reduce the reluctance and increasing the length of the magnetic circuit will increase the reluctance.

PERMEABILITY

Silver, copper, and aluminum have different conductivities. Silver has the highest conductivity and is the best conductor, then copper and then aluminum. A copper wire has less resistance than an aluminum wire of the same size and length.

Similarly, different magnetic materials have different permeabilities. The permeability of the core material determines the total reluctance of the coil; when the permeability goes up, the reluctance goes down and vice versa.

The permeability of air and all other nonmagnetic materials is considered to have the numerical value of 1. Magnetic materials all have higher permeability values than 1, ranging from about 50 all the way up to 10,000 or even higher for certain special alloys. Thus if the permeability of a material is 10, we can expect 10 times the magnetic flux through this material than we would have through air for the same number of ampere turns.

MAGNETIC FLUX

You already know that magnetic flux in a magnetic circuit corresponds to current in an electric circuit. In an electric circuit, current is equal to the voltage in volts divided by the resistance in ohms. In a magnetic circuit the flux is equal to the magnetomotive force divided by the reluctance.

In practical magnetic circuits you will not have to calculate the magnetic flux. Even if you performed this calculation it would be of no value to you. However, it is important that you understand what magnetomotive force, reluctance, and magnetic flux are, and how they are related to each other.

You can increase the amount of flux in a magnetic circuit either by increasing the magnetomotive force or by decreasing the reluctance. You can decrease the amount of flux either by decreasing the magnetomotive force or by increasing the reluctance. Every change in flux is thus due to a change in either magnetomotive force or to a change in reluctance.

SUMMARY

Magnetic circuits are like electrical circuits in many ways. In a magnetic circuit there is a force which is the equivalent of voltage in an electrical circuit. We call this force a magnetomotive force. This force produces flux or magnetic lines of force which travel around the magnetic circuit. The flux or magnetic lines of force have some opposition; the opposition is known as reluctance. The magnetic lines of force are roughly the equivalent of the current in an electrical circuit and the reluctance is the equivalent of the resistance. In a circuit with a given magnetomotive force, if the reluctance is lowered, the flux in the circuit will increase. On the other hand, if the reluctance increases, the flux decreases.

Remember that the unit of magnetomotive force is an ampere-turn. A current of 1 ampere flowing through a coil of one turn produces a magnetomotive force of 1 ampere-turn. If we double the number of turns so that we have two turns and the current remains the same, the magnetomotive force produced will be two ampere-turns.

Remember the term permeability. The permeability of a material indicates the ability of the material to pass magnetic lines of force. The permeability of air is 1. Other non-magnetic materials such as paper, ceramic, glass etc. also have a permeability of 1 - in other words they act just like air insofar as a coil is concerned. However, magnetic materials such as iron and various alloys of magnetic materials have a much higher permeability than air. When we speak of a coil with a high permeability core we are talking about a coil with a core that has a much higher ability to pass magnetic lines of force than air.

Make sure that you remember these important terms used in conjunction with magnetic circuits. You will run into them many times in the future, and if you understand them now, then you'll understand the way in which they are used later. If you have any doubts about this section of this lesson be sure to review before going on. Once you are sure you have mastered the material in the lesson, do the self-test questions. Again, if you have trouble with any of the self-test questions, don't hesitate to go back to the text and restudy it. The purpose of the self-test questions is to help you be sure you have mastered the important points in the section of the lesson. If there is any self-test question you are not sure of, this indicates that you need to go back and spend some extra time on this section of the lesson before going ahead.

SELF-TEST QUESTIONS

(a) Fill in the missing word: Magnetic lines of force form ____ loops.

(b) Where will most of the lines of force produced by a coil be concentrated?

(c) What type of coil is a coil wound on a cardboard form?

(d) What are laminations?

(e) What purpose does an iron-core serve in a coil?

(f) What is magnetomotive force?

(g) What is the unit of magnetomotive force?

(h) If the current flowing through a 25-turn coil is 2 amperes, how many ampere-turns are produced?

(i) What is reluctance?

(j) Does the magnetic circuit in an air-core coil have a higher reluctance than a magnetic circuit in an iron-core coil?

(k) Which has the lowest reluctance, paper, air, aluminum or copper?

(l) What is permeability?

(m) What is the relation between flux, magnetomotive force and reluctance?

Using Coils To Produce Voltage

In most of the electronics applications of coils with which you are concerned, coils will be used to produce a voltage. An obvious example of this application is the transformer. In a transformer, a voltage is applied to one winding and this voltage causes a current to flow through this winding. This sets up a magnetic field, which in turn induces a new and completely separate voltage in another coil. Here two coils have been used to produce a voltage.

Another type of device where a coil is used to produce voltage makes use of a coil placed in the field of a permanent magnet to take a wave or signal other than an electrical signal and produce an electrical signal from it. Such a device is called a transducer. An example of a transducer using this principle is a dynamic microphone. A dynamic microphone uses coils placed in a magnetic field to convert an audio signal, which is actually a wave or vibration in air, to an electrical signal, which is the electrical equivalent of sound.

If you understand how a voltage can be produced by a coil, you will have mastered the most important point in understanding how coils work. To see how a voltage can be induced in a coil, we must first learn something about flux and flux linkages and then see how changing the flux linkages of a coil will produce a voltage in a coil.

FLUX LINKAGES

Let us see what we mean by flux linkage. Suppose we have a magnet that produces a single magnetic line of force. If the magnet is brought near a coil having one turn such as shown in Fig. 10A, we will have one magnetic line linking with or passing through a one-turn coil, and we will have one flux linkage. If we had ten turns on the coil and the one magnetic line of force passed through all ten turns, then we would have ten flux linkages as shown in Fig. 10B. However, if the single magnetic line passed through and linked only six turns of the 10-turn coil, as shown in Fig. 10C, we would have only six flux linkages. Thus the term

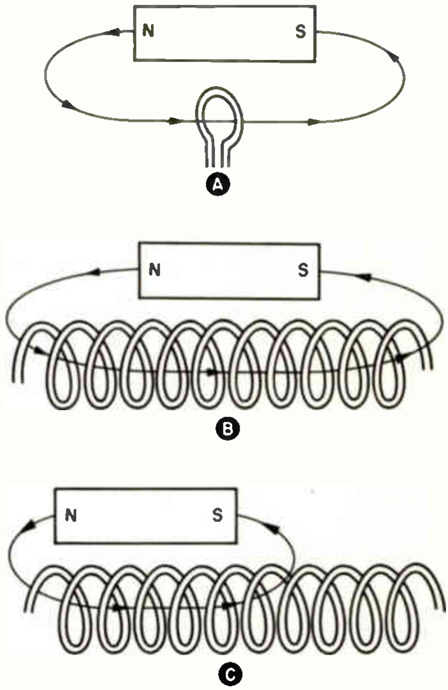


Fig. 10. How the number of flux linkages can be changed. In A there is one flux linkage; in B, ten; and in C, six.

"flux linkage" is an indication of the number of magnetic lines of force passing through and linking the turns on the coil. If we have a magnet that produces 100 magnetic lines, and the entire 100 lines linked to a coil having 80 turns, the number of flux linkages would be 80 times 100, or 8000 flux linkages.

Changing Flux Linkages.

Now let's look at Fig. 11A. Here we have a magnet with 10 magnetic lines of force, but actually only two of them are cutting through a coil with 5 turns on it, so we have a total of 10 flux linkages. As you can see, part of the flux is lost--it does not cut through the coil. This is called leakage flux or flux leakage. If we suddenly move the magnet to the position shown in Fig. 11B so that the magnet is placed inside the coil and all ten lines cut through the five turns of the coil, we have a total of 50 flux linkages. When the number of flux linkages increases from 10 to 50 there will be a voltage induced in the coil. This voltage is known as an induced voltage.

If the magnet is then moved away from the coil so that the number of flux linkages is changed from 50 back to 10, we will again have a voltage induced in the coil.

In each of the two examples given, we had a change of 40 flux linkages. If we had a stronger magnet so that the number of lines of force was greater, and therefore the change in flux linkages was greater, we would have a greater voltage induced in the coil.

In moving the magnet either towards or away from the coil, the voltage that will be induced in the coil will depend upon the speed with which the magnet is moved. If the magnet is moved slowly so that the number of flux linkages changes slowly, the voltage induced in the coil will be small. However, if the magnet is moved rapidly so that the change in flux linkage occurs very quickly, the voltage induced in the coil will be higher. If it took one second to move the magnet so that the number of flux linkages changed from 10 to 50, we would get a certain voltage induced in the coil. The exact value is not important to this discussion. However, if we were to move the magnet so that the number of flux linkages was changed from 10 to 50 in 1/100th of a second, we would get exactly 100 times as much voltage as before. The faster the rate of change in flux linkages, the greater the induced voltage will be.

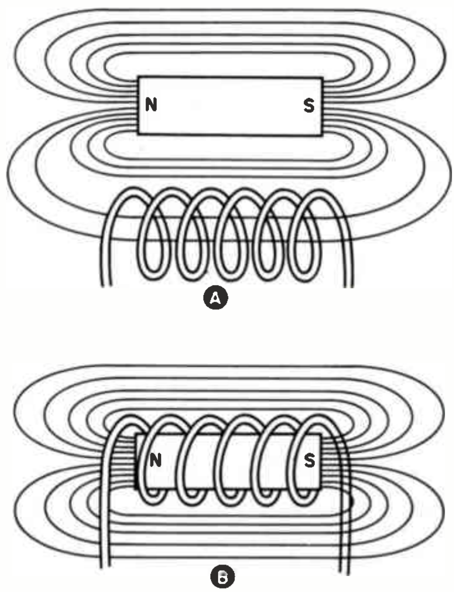


Fig. 11. In A we have 10 flux linkages; in B we have 50 when the same magnet is moved inside the coil.

LENZ'S LAW FOR COILS

The voltage induced in a coil always acts in a definite direction. In other words, the voltage has a definite polarity. This polarity at any given instant depends on just two things--on the direction of the original flux, and on whether the flux linkages are increasing or decreasing.

The exact relationship between these things is expressed by a famous electrical law known as Lenz's Law. The law is named after the man who was the first to realize that the direction in which an induced voltage will act can always be predicted before it is produced.

When the number of flux linkages cutting a coil is changed, a voltage will be induced in the coil. This induced voltage will have a polarity such that if the circuit is complete, it will send a current through the coil which opposes the change in magnetic flux. In other words, if the flux linkages are increasing, the induced voltage will tend to send a current through the coil that would produce a magnetic flux which would oppose the original coil flux to try to keep it from increasing. On the other hand, if the flux linkages are decreasing, the induced voltage will be of such a polarity that it will produce a current which in turn will produce a flux which aids the original flux and tends to prevent the flux from decreasing.

This is an extremely important law and can be better understood by referring to the circuits shown in Fig. 12. In Fig. 12A, we have a magnetic circuit with two flux lines cutting through a 5-turn coil, which gives us ten flux linkages. As the magnet is moved away from the coil, reducing the number of flux linkages as shown in Fig. 12B, a voltage will be induced in the coil, and current will flow through the coil. Current flowing through the coil will set up a magnetic field which will aid the flux linkages already existing. As long as the number of flux lines is changing, the induced voltage will be present and will cause the induced current to produce flux lines as shown.

If, as shown in Fig. 12C, the magnet is moved back into the coil, the number of flux linkages would tend to increase. However, a voltage will be induced in the coil that will cause

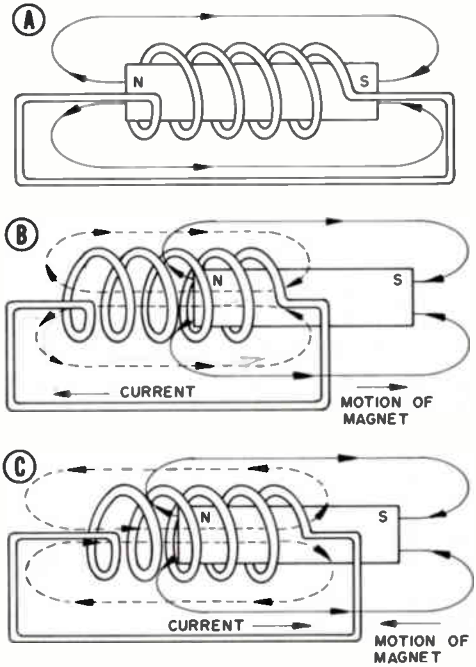


Fig. 12. Changing the number of flux linkages induces a voltage in the coil. The polarity of the voltage depends upon whether the number of flux linkages is increasing or decreasing.

a current to flow in the opposite direction and set up its own lines of force to oppose the flux lines from the magnet. In other words, if there is any change in the number of flux linkages through the coil, a voltage is induced in that coil that will cause a current to flow which in turn produces its own flux to oppose the change in flux linkages.

If the number of flux linkages is decreasing, the induced voltage will have a polarity such that it will cause a current to flow to oppose this decrease in flux linkages; and on the other hand, if the flux linkages are increasing, then the induced voltage will have a polarity that will cause a current to flow to oppose this increase in flux linkages.

METHODS OF CHANGING FLUX LINKAGES

There are three methods of producing changes in the flux linkages in a coil. They are: by cutting through magnetic lines of force; by changing the reluctance; and by changing the current flowing in the coil.

Cutting Lines of Force.

You have already seen an example of this method of producing changes in flux linkages when you studied generators in an earlier lesson. You learned that when a conductor is moved through a magnetic field it cuts the magnetic lines of force, and a voltage is induced in the conductor. We get this induced voltage because the motion of the conductor changes the flux linkages as the conductor passes through the magnetic lines of force.

In a generator, instead of moving a single wire through a magnetic field, a coil is rotated in the magnetic field. As the coil is rotated, it moves through and cuts through the magnetic lines of force produced by a permanent magnet or an electromagnet, and a voltage is induced in the coil. The voltage induced in the coil will have a polarity such that the current that will flow when the coil is connected to an external circuit will set up a magnetic field in the coil which opposes the change in flux linkages producing the voltage.

Changing the Reluctance.

Any change in the reluctance of a magnetic circuit will change the amount of flux which passes through the coil, thus changing the flux linkages through the coil and inducing a voltage. Remember, whenever there is a change in flux linkages, a voltage is induced.

An example of this method of producing a voltage is the variable reluctance phono pickup used in many record players. A simplified drawing of one is shown in Fig. 13. The needle, or stylus, is mounted on a cantilever spring, which moves between two coils. The other end of the cantilever spring is connected to the south pole of a permanent magnet. A T-shaped yoke connects the other end of the magnet to two pole

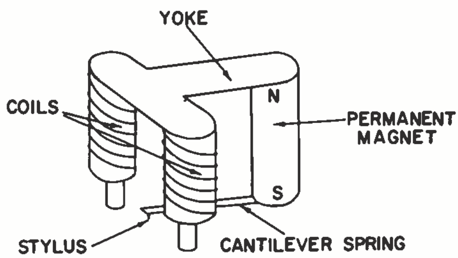


Fig. 13. A variable reluctance phono pickup.

pieces on which two coils are wound. The flux path goes from the magnet through the yoke, and the two pole pieces, across the air gap to the cantilever spring, and through it back to the magnet. As the needle follows the record grooves, it moves from side to side, nearer one or the other of the two coils. As it does so, the air gap on one side decreases, so the reluctance on that side decreases, and the flux increases. At the same time, the air gap on the other side becomes wider, increasing the reluctance and decreasing the flux. Since the flux changes are in opposite directions, the voltages induced in the two coils will be of the opposite polarity. The two coils are connected in such a way that the two voltages are added in the output. The change in flux linkages will induce a voltage in the coil.

Changing the Coil Current.

When two coils are arranged as shown in Fig. 14, the flux produced by coil L1 passes through coil L2. As long as the current through L1 remains constant, the flux produced by this coil will remain constant, and there will be no change in the flux linkages in L2. However, if the current is changed by changing the setting of the rheostat, there will be a change in the flux produced by L1 and hence a change in the number of flux linkages in L2. This will induce a voltage in L2.

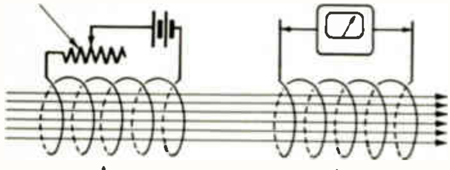
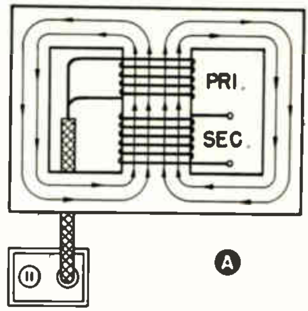


Fig. 14. When two coils are arranged as shown above, the flux produced by one passes through the other.



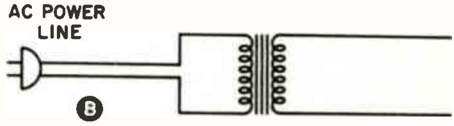


Fig. 15. A power transformer. The primary winding is connected directly to the power line. As the current through the primary working varies, the flux produced by L1 will vary, resulting in a change in flux linkages through the secondary winding, inducing a voltage in the secondary.

Of course, this is not a practical way of inducing a voltage in L2 because the rheostat setting would have to be changed continually and at a rapid rate in order to produce any appreciable voltage in L2. A more practical application would be to apply an ac voltage to L1 in place of the battery and the rheostat.

Since the ac voltage is continually changing, this means that the flux is constantly changing, which in turn will result in there being a voltage continually induced in L2.

A practical application of this principle is in the power transformer as shown in Fig. 15. We have already mentioned the transformer several times before and you are aware that a transformer is simply two coils wound on a common core. In the transformer shown in Fig. 15, the one winding called the primary is connected directly to the ac power line. The ac voltage which is a varying voltage will cause a varying current to flow through the primary winding. As the current varies, the flux produced by the primary will change, resulting in a change in the flux linkages cutting the secondary winding. This change in flux linkages will induce a voltage in the secondary winding. As we pointed out before, whether this voltage is higher or lower than the primary voltage will depend upon whether the secondary winding of the transformer has more or fewer turns than the primary winding.

SUMMARY

There are several important facts that you should remember from this section of the lesson. First, remember that a voltage is induced in a coil when the number of flux linkages changes. Either an increase in the number of flux linkages or a decrease in the number of flux linkages will induce a voltage in the coil.

Lenz's Law of coils is important. It states that the induced voltage always acts in such a direction that it tends to oppose the original change in flux linkages.

Changes in flux linkages can be produced by cutting through magnetic lines of force, by changing the reluctance in the magnetic circuit or by changing the current flowing through the circuit.

In the next section of this lesson you will learn more about coils. You will learn how the electrical characteristics of coils are expressed and also see why the opposition that a coil offers to the flow of ac through it is much higher than it is to dc. However, before going ahead with the next section of the lesson, it is important that you understand the material covered in this section. Therefore you should review this section if necessary and then answer the self-test questions. Be sure you are able to answer all of the following self-test questions before you go ahead with the next section of the lesson.

SELF-TEST QUESTIONS

(n) What is meant by flux linkages?

(o) If three magnetic lines of flux cut through four turns of a coil, how many flux linkages have we?

(p) If one hundred flux linkages cut through a coil, and this number of flux linkages does not change, what will the voltage induced in the coil be?

(q) According to Lenz's Law, if a change in the number of flux linkages cutting a coil occurs, will the voltage induced in the coil produce a current, which in turn will build up a magnetic flux that will aid or oppose the original change in flux linkages?

(r) When the number of flux linkages cutting a coil is reduced, will the field produced by the induced voltage aid or oppose the original lines of force?

(s) Name three methods of changing flux linkages.

(t) Give a practical example where changing the reluctance in a magnetic circuit is used to produce a voltage.

(u) Give a practical example of where changing the coil current produces a voltage.

Inductance

In the preceding section of this lesson you learned that if the number of flux linkages cutting the turns of a coil changes, there will be a voltage induced in the coil. The exact amount of voltage that will be induced in the coil will depend upon how great the change in flux linkages is, and how rapidly it occurs. It will also depend upon the coil itself. The property of the coil that will govern or determine the voltage induced in the coil is called inductance. The inductance of a coil will depend upon the number of turns of wire on the coil and upon the permeability of the core material. Saying that a coil has inductance is just about the same as saying a resistor has resistance. Inductance is a basic property of a coil and it indicates how much voltage will be induced in the coil for a given change in flux linkages. Before we go further with this idea of inductance, you should learn something about self-induced voltages.

SELF-INDUCED VOLTAGES

When a coil is brought near a magnetic field and the strength of the field is suddenly changed, there will be a change in the number of flux linkages through the coil. You know that this will result in a voltage being induced in the coil.

However, now let us consider a coil that is completely removed from any external magnetic field. If a voltage source is connected to the coil, current will flow through the coil and this current will set up a magnetic field. The magnetic field produced by the coil will produce lines of flux. These lines of flux will link through the turns of the coil as shown in Fig. 16.

If the current flowing through the coil is suddenly changed, the strength of the magnetic field will change and this will result in there being a change in the number of flux linkages passing through the turns of the coil. This will have exactly the same effect as changing the flux

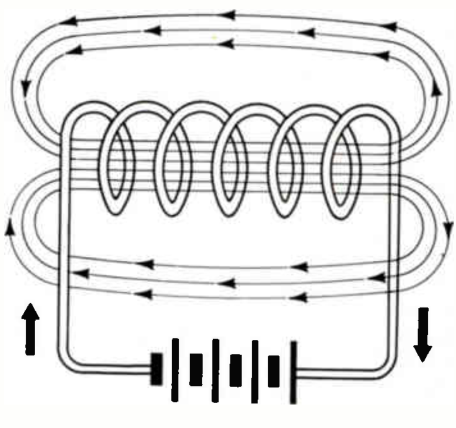


Fig. 16. If a voltage source is connected to a coil, current will flow through the coil and a magnetic field will be set up.

linkages produced by an external magnet will have. There will be a voltage induced in the coil and this voltage will be such that it will tend to oppose the change produced. In other words, the induced voltage will cause a current to flow in such a direction as to produce a magnetic field which tends to oppose the change in the magnetic field. The voltage induced in this manner is known as self-induced voltage.

It is important for you to realize that an induced voltage in a coil always opposes the change producing it. For example, in the circuit shown in Fig. 16, if the voltage is reduced to lower the current and hence the number of flux linkages, the self-induced voltage induced in the coil will have a polarity that both aids the applied voltage and tries to keep the current constant so that the number of flux linkages will not change. On the other hand, if the applied voltage is suddenly increased, then the voltage induced in the coil will oppose the applied voltage to, once again, try to keep the current flowing through the coil constant and hence the number of flux lines constant.

The induced voltage in a coil obeys Lenz's Law. Its polarity is such that it tries to oppose the change that produced it.

UNITS OF INDUCTANCE

For a given change in flux linkages, the voltage that will be induced in a coil will depend upon the inductance of the coil. The unit of inductance is the henry. It is named after Joseph Henry, an outstanding scientist who did a great deal of experimenting with coils.

There are a number of scientific ways of defining the henry, but these are of no importance to the electronics technician. One simple definition of the henry that can be used is as follows: if the voltage induced in a coil is 1 volt when the strength of the current flowing through the coil changes at a rate of 1 ampere per second, the coil has an inductance of 1 henry. In other words, if a coil has an inductance of 1 henry, a current change of 1 ampere per second will induce a voltage of 1 volt in the coil.

Large iron-core coils frequently have quite high inductances. You will find iron-core coils in electronic equipment having inductances of 20 or 30 henrys. In some cases you may find iron-core coils having inductances ranging as high as 1000 henrys.

Most air-core coils have a very small inductance. For convenience in specifying inductance values of air-core coils and some small iron-core coils, two other units are used, the millihenry and the microhenry. Just as the milliampere is one thousandth of an ampere, the millihenry is one thousandth of a henry; just as a microampere is one millionth of an ampere, the microhenry is one millionth of a henry. The unit millihenry is usually abbreviated mh and the microhenry is abbreviated μh.

To convert from henrys to millihenrys or microhenrys you use exactly the same procedure as in converting amperes to milliamperes or microamperes. To convert henrys to millihenrys, you multiply by 1000 or simply move the decimal point three places to the right. To convert henrys to microhenrys you multiply by 1,000,000 or move the decimal point six places to the right. To convert millihenrys to microhenrys, you multiply by 1000 or move the decimal point three places to the right.

To convert from microhenrys to henrys, you divide by 1,000,000, or move the decimal point six places to the left. To convert from millihenrys to henrys, you divide by 1000 or move the decimal point three places to the left. To convert from microhenrys to millihenrys you divide by 1000 which is the same as moving the decimal point three places to the left. If you have no difficulty converting from amperes to milliamperes and microamperes and back again to amperes you should have no problem converting between henrys, millihenrys and microhenrys.

FACTORS AFFECTING INDUCTANCE

There are several factors that affect the inductance of a coil. As you might expect, one of the chief factors is the number of turns on the coil. You can expect a coil having 200 turns to have a higher inductance than a coil having 100 turns would on the same type of core.

The inductance of a coil is also affected by the shape and size of the coil. As an example, an air-core coil wound on a round form six inches in diameter will have a higher inductance than a coil with the same number of turns wound on a form one inch in diameter. In the coil with the smaller diameter, many of the lines of flux will escape or cut through only a few turns of the coil; in other words there will be considerable flux leakage. On the other hand, in the larger coil more flux lines will cut through each turn of the coil, resulting in a greater number of flux linkages, which in turn will give the coil a greater inductance.

The inductance of a coil is affected by the core material. If a magnetic material is placed in the core of a coil, the magnetic path will have a much lower reluctance; there will be more flux and a much greater number of flux linkages than there would be in a similar coil without an iron-core. The exact material placed inside the coil also affects the inductance. The higher the permeability of the core material, the greater the inductance of the coil will be.

Before leaving this section of the lesson it should be pointed out that inductance is not limited to coils alone. Even a straight wire has some inductance, because when a current flows through the wire, a magnetic field is set up around the wire and the wire will be cut by magnetic lines. Of course, the inductance of a straight wire is much lower than it would be if the wire were wound into a coil, but nevertheless every piece of wire does have inductance. In most cases this inductance is so low it has no effect on the circuit performance, but in some ultra-high-frequency electronic equipment straight wires or tubing are actually used as "coils."

Because the most important property of a coil is its inductance, electronics men often call coils "inductors" or "inductances." The term inductance not only includes coils, but in the case of ultra-high-frequency equipment may include a straight piece of tubing that is to be used as the inductance in one of the circuits.

INDUCTIVE REACTANCE

You have learned that when a coil is connected to a voltage source as in Fig. 16, and the voltage is changed, there is a voltage induced in the coil that opposes the change in voltage. Now let us consider what happens when a coil is connected to an ac voltage source as in Fig. 17. Here the voltage is continually changing.

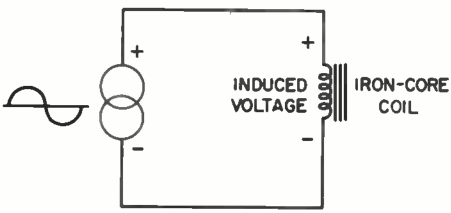


Fig. 17. A coil connected to an ac voltage source.

During the first quarter cycle when the ac voltage is increasing and has the polarity shown in Fig. 17, the polarity of the induced voltage will be as shown so it will oppose the ac voltage as it tries to increase. Thus, the induced voltage acts to oppose and limit the change in current in the coil. If there were no voltage induced in the coil, the current would increase as the voltage increased, and the actual value of current flowing at any instant would depend only on the voltage applied and the dc resistance of the coil. However, since the induced voltage is of the opposite polarity to the applied voltage, the induced voltage has the effect of opposing the applied voltage and limiting the current change in the coil. This self-induced voltage is known as "counter" or "back" emf (electromotive force). It is the induced ac voltage that appears across the coil. This ac voltage drop is just the same as the voltage drop across a resistor caused by current flowing through a resistor. In other words, the counter emf is the ac voltage drop across a coil caused by the opposition that the coil offers to the alternating current. This opposition that the coil offers is not the same as resistance because it affects only ac, and not dc. The opposition is known as inductive reactance and it is measured in ohms.

The amount of voltage induced in a coil will depend upon how rapidly the change in flux linkages occurs. When an ac voltage is applied to a coil, the speed with which the number of flux linkages changes depends upon the frequency of the ac voltage. Thus the change in flux linkages occurs more rapidly if the frequency is 100 cycles than it would if the frequency were only 10 cycles. Therefore an ac current with a frequency of 100 cycles flowing through a coil will induce more voltage than a current of the same strength but a frequency of only 10 cycles. This means that the inductive reactance depends upon the frequency.

The inductive reactance of a coil can be determined by multiplying the inductance of the coil in henrys times 6.28 times the frequency in cycles.

Electronics technicians use symbols to provide a short convenient way of expressing this relationship. The symbol used for reactance is X. A small capital letter L following the X and written X_L is used to indicate inductive reactance. The letter f is used to represent frequency, and the letter L is used to represent inductance. Thus the expression for inductive reactance of a coil in ohms can be written:

$$X_L = 6.28 \times f \times L$$

The number 6.28 is two pi. Pi is the Greek letter π (pronounced pie) which represents the number 3.14. You have probably seen this number before--it is used to find the area of a circle. Remember that the area of a circle is π times the radius squared. 2π is 6.28, which is a number that appears in many electrical formulas. Sometimes you will see the expression for inductive reactance written:

$$X_L = 2\pi fL$$

Now let's see how we use this formula to find the inductive reactance of a coil.

Example 1: Suppose we want to know the inductive reactance of a 50-henry choke at 100 cycles. The formula is:

$$X_L = 6.28 \times f \times L$$

Substituting 100 for f and 50 for L gives:

$$X_L = 6.28 \times 100 \times 50$$

Multiplying these numbers gives us 31,400 ohms. This is the inductive reactance of the coil at a frequency of 100 cycles. At a frequency of 50 cycles, the inductive reactance would be half this figure, and at a frequency of 200 cycles per second, the inductive reactance would be twice this figure. We say that the inductive reactance of a coil varies directly as the frequency varies. If the frequency increases, the reactance increases, and if the frequency decreases, the reactance also decreases.

Example 2: Suppose we want to know the inductive reactance of a 10-henry choke at a frequency of 100 cycles. Substituting 100 for f and 10 for L gives us:

$$X_L = 6.28 \times 100 \times 10$$

Multiplying this gives us:

$$X_L = 6280 \text{ ohms.}$$

Notice that this is less than in the case of the 50-henry coil. Thus the inductive reactance also varies directly as the inductance of the coil varies. Reducing the inductance reduces the reactance and increasing the inductance increases the reactance.

As an electronics technician you will seldom have to work on a problem of this type. However, it is important for you to remember that the inductive reactance of a coil varies directly both with the frequency and with the inductance of a coil.

You might wonder what the inductive reactance is of a small coil consisting of only a few turns. At low frequencies of a few hundred cycles, the reactance is so low that in most cases it can be ignored. However, when small coils are used in high-frequency circuits, their inductive reactance can be appreciable. Let's take as an example a 10-microhenry coil used at a frequency of 100 megacycles.

10 microhenrys is .000010 henry, and 100 megacycles (abbreviated mc) is 100,000,000 cycles.

$$X_L = 6.28 \times f \times L$$
$$X_L = 6.28 \times 100{,}000{,}000 \times .000010 = 6280 \text{ ohms.}$$

Thus, even though the inductance of the coil is quite small, at the frequency of 100 mc it has as high an inductive reactance as the 10-henry coil had at 100 cycles.

We mentioned previously that inductive reactance is the opposition a coil offers to the flow of ac current through it. A coil has no inductive reactance to dc. You can see that this must be true from the formula for inductive reactance. The frequency of dc is zero and so if we substitute zero in the formula for inductive reactance, then we have $6.28 \times 0 \times L$. Whenever you multiply anything by 0, no matter how large it is, the result is 0, so the inductive reactance is 0. The only opposition a coil will offer to the flow of dc through it is due to the resistance of the wire used to wind the coil. The wire will have a certain resistance, and this resistance will oppose the flow of dc through the coil in just the same way as it would if it were one long piece of wire and we tried to pass dc through it. The ac reactance of a coil, on the other hand, is an entirely different thing; it is the opposition the coil offers to the flow of ac through it due to the inductance of the coil, and it will be much higher than the dc resistance of the coil.

MUTUAL INDUCTANCE

When two coils are placed near each other so that some of the flux produced by one coil will cut through the turns of the other coil, the coils are said to be mutually-coupled through their magnetic fields. The coils might actually be wound on the same iron core or they might simply be placed near each other. When coils are placed so they are mutually coupled, any change in the flux in one coil will induce a voltage in the other coil.

Mutual inductance is measured in henrys, just as the inductance of a single coil. The mutual inductance is usually represented in formulas by the letter M. The greater the value of mutual inductance, the greater will be the voltage in one coil when the current through the other changes.

Mutual inductance is defined in the same way as inductance - when a primary current changes at a rate of 1 ampere per second, if the voltage induced in the secondary coil is 1 volt, the mutual inductance is 1 henry.

Mutual inductance depends upon the size of both coils, the number of turns of each coil and how many flux linkages from one coil cut the turns of the other coil.

COILS IN SERIES AND PARALLEL

When we consider coils connected in series or in parallel, there are two different cases to consider. The first and simplest is if the coils are located some distance from each other so that their magnetic fields do not affect each other. When the coils are connected in series as shown in Fig. 18A, the combined inductance is the sum of the individual inductances. In other words, the total inductance is obtained simply by adding the inductances of the individual coils. This should be easy to remember because in this respect coils are like resistors.

When coils are connected in parallel as shown in Fig. 18B, the total inductance will be less than the inductance of the smallest coil in the

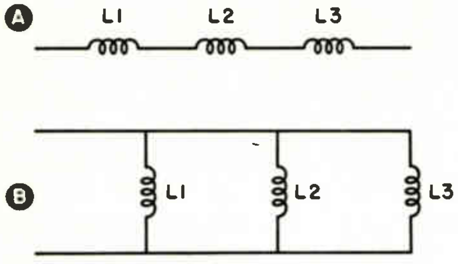


Fig. 18. Coils connected in series (A); coils connected in parallel (B).

group. Again, this is just like resistors connected in parallel - remember that when resistors are connected in parallel, the total resistance is always less than the resistance of the smallest resistor.

It is easy to see why coils connected in parallel act this way. For example, looking at Fig. 18B, if L1 has an inductive reactance of 100 ohms and is connected across a 100-volt source, an ac current of 1 amp would flow through the coil. If the second coil of equal inductance is connected in parallel with it, an ac current of 1 amp will also flow through it. Now we have a current of 2 amps flowing in the circuit and therefore, the opposition or inductive reactance must have decreased. This means that the total inductance of the two coils in parallel must have decreased. In fact, with two equal coils, the effective inductance of the two in parallel will be equal to one half the inductance of either coil.

When coils connected in series are placed close together so that some mutual inductance exists, there is interaction between the coils, and the combined inductance can no longer be figured simply by adding the inductances of the individual coils. In this situation, we must consider the mutual inductance in the circuit and also how the coils are connected together.

Let us look at the first case, where the two coils are connected in series so that the flux from one coil aids the flux from the other. In other words, the magnetic lines are flowing in the same direction. Here if the inductance of the two coils is represented by L1 and L2 and the mutual inductance by M, the total inductance (L_T) of the two coils connected in series will be equal to:

$$L_T = L1 + L2 + 2M$$

If the connections to one of the coils are reversed, its magnetic field will oppose the magnetic field of the other. Under these circumstances the total inductance of the two coils connected in series will be:

$$L_T = L1 + L2 - 2M$$

SUMMARY

You have studied a great deal of important material about coils in this section and it would be worthwhile to read the section over several times to be sure that you have understood everything covered.

You have learned that the electrical property that describes coils is called inductance, and that inductance is measured in henrys. A coil has an inductance of 1 henry when a current change of 1 amp per second induces a voltage of 1 volt in the coil.

You learned that when the current flowing through a coil changes, there is voltage induced in the coil that opposes the change that produces it. This voltage is a self-induced voltage and is called counter emf or back emf.

Coils have a property called inductive reactance. Inductive reactance is the opposition that a coil offers to the flow of ac through it. Inductive reactance is measured in ohms and is somewhat similar to resistance inasmuch as it opposes the flow of ac through the coil.

When two coils are placed near each other, the flux lines of one coil will cut through the other coil, and the coils are said to be mutually-coupled. The amount of coupling is determined by the nearness of the coils to each other and by the shape and size of the coils. This coupling is called mutual inductance. The mutual inductance of two coils is measured in henrys.

When coils are connected in series, the total inductance is equal to the sum of the individual inductances, and when they are in parallel, the total inductance is less than the inductance of the smallest coil. When mutually-coupled coils are connected in series, the total inductance is $L1 + L2 + 2M$ when the magnetic field of the two coils aid each other, and $L1 + L2 - 2M$ when the magnetic fields oppose each other.

SELF-TEST QUESTIONS

(v) What is the name given to the property of a coil which will determine the voltage induced in it?

(w) What is a self-induced voltage?

(x) If the voltage applied to a coil is suddenly increased, will the self-induced voltage produced in the coil aid or oppose the applied voltage?

(y) What is the unit of inductance?

(z) How is the unit of inductance defined?

(aa) Name three factors which affect the inductance of a coil.

(ab) What is the inductive reactance of a coil?

(ac) What is the unit in which the inductive reactance of a coil is measured?

(ad) What is the inductive reactance of a 10-henry coil at a frequency of 60 cycles?

(ae) What is meant by mutual inductance?

(af) If two coils, one having an inductance of 6 henrys and the other having an inductance of 8 henrys are placed some distance apart so that there is no mutual inductance between them, what will the total inductance of the two coils be if they are connected in series?

(ag) Two coils, one having an inductance of 4 henrys and the other having an inductance of 3 henrys have a mutual inductance of 2 henrys. If the coils are connected in series aiding, what will the total inductance be?

(ah) If an 8-henry coil and a 7-henry coil that have a mutual inductance of 3 henrys are connected in series opposing, what will the total inductance of the two coils be?

(ai) Convert 2.2 henrys to millihenrys.

(aj) Convert 100 microhenrys to henrys.

Ohm's Law for Coils

You will remember from earlier lessons that the current that will flow in a circuit depends upon the voltage applied and upon the resistance of the circuit. This rule can be applied to ac circuits as well as dc circuits, but in an ac circuit, you substitute the total opposition offered to the flow of current for the resistance. In an ac circuit, the total opposition to current flow is called impedance and is represented by the letter Z. In this section of this lesson, you will study Ohm's Law for coils, you will learn what impedance is and how to find impedance in ac circuits, and you will learn about another important thing in ac circuits which is called phase.

PHASE

Before you can understand why impedance is important in ac circuits you must understand what we mean by phase. Phase is important. It is something that you will run into all the way through your study in electronic circuits. Time and time again you will see the expressions, "in phase", "out of phase" and "phase shift." Since phase is so important, learning what it is now will simplify your studies later.

In a circuit made up only of resistance, if we increase the voltage, the current will increase immediately. Changes in current can be produced instantly by changing the voltage in the circuit. In other words, the current follows the voltage changes instantly. If the voltage increases, the current increases instantly; if the voltage decreases, the current decreases instantly. The current is in phase with the voltage. This idea simply means the change in the voltage will produce the same change in the current flowing in the circuit.

This is not true of circuits containing coils. If you have a constant dc voltage connected across a coil, the current that flows depends only on the dc resistance of the coil; in other words, the resistance of the wire used to wind the coil. If you suddenly increase the voltage applied to the coil, there is immediately a change in the number of flux linkages cutting the various turns of the coil. This induces a voltage in the coil, and the induced voltage opposes the change in applied voltage. When the current changes because of the increase in applied voltage, it will be limited by the induced voltage as well as the resistance. The inductance of the coil opposes the change in current through the coil - this effect is called inductive reactance. Gradually the current will increase from the value that it was originally if the voltage were increased. As the current increases, it finally reaches its new value; the self-induced voltage in the coil decreases until finally when the current becomes constant, the self-induced voltage in the coil will disappear.

Now let us look at the circuit shown in Fig. 19. Here we have a resistor and a coil connected in series and the two are connected across an ac generator. We are

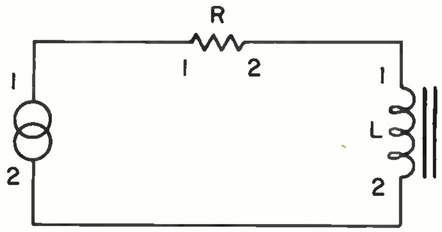


Fig. 19. A coil and resistor connected in series across an ac source.

going to examine the ac current flowing through the circuit to see what happens to the voltage across the resistor and the voltage across the coil, as the current goes through its cycle.

In Fig. 20A we have shown a single current cycle from the generator in

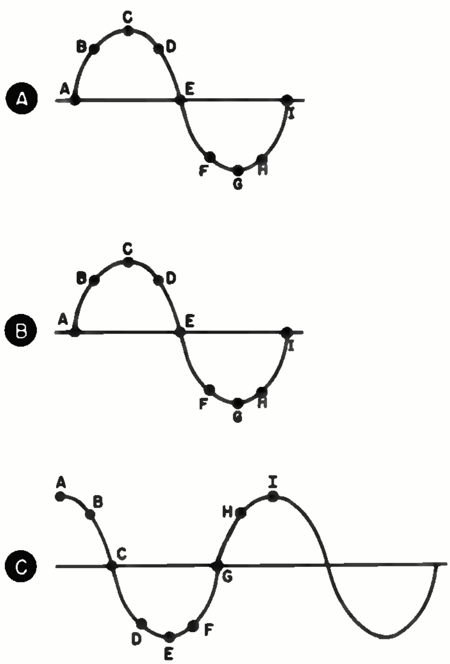


Fig. 20. Generator current is shown at A, resistor voltage at B and coil voltage at C.

Fig. 19. At the present, we are not worried about the voltage across the generator, we are simply concerned with the current. Remember that the ac voltages and current waveforms from the generator will be sine waves. The start of the sine wave cycle is marked point A. Let's consider that at this instant current is just starting to flow from terminal 1 of the generator around the circuit towards terminal 1 of the resistor. At the instant the cycle starts at terminal A, the current is increasing at its maximum rate of change. Notice that as the cycle moves from point A to point B, the curve is flattening out until finally at point C the current is neither decreasing or increasing, it is at a constant value for just an instant. From point C to point D the current begins to decrease. The rate at which it decreases is increasing from C to D and it continues to change at an even more rapid rate until it reaches point E at which instant the current is at 0 for just a moment. However, even though the current is at 0 for an instant at point E, the rate of change is very rapid; the instant before it reaches point E it is flowing in one direction, exactly at point E it drops to 0 and then at the instant it passes point E and starts towards point F, it begins to flow in the opposite direction. As the cycle moves from E to F, the rate at which the current is changing begins to decrease until finally at point G, once again while the current is flowing at its maximum value, the rate at which it is changing for just an instant at point G drops to 0. From point G to point H, the current again starts to drop to 0 and the rate at which it is dropping to 0 begins to

increase from point G to point H and continues to increase until it reaches its maximum rate of change at point I.

Now let us see what happens as this current cycle flows around the circuit. Let us consider the voltage across the resistor R first. As the current flows from terminal 1 to terminal 2 of resistor R, the voltage that will be produced across the resistor will depend upon the current flowing through the resistor. As the current wave builds up from point A to point B and then to point C as shown in Fig. 20A, a voltage wave will be built up as shown in Fig. 20B. Point A represents 0 voltage and point C represents maximum voltage. Maximum voltage at point C on the curve B will be reached at exactly the same instant as the maximum current flow is reached at point C on curve A. As the current through the resistor begins to decrease and finally reaches 0 at point E on curve A, the voltage across the resistor will follow curve B reaching 0 at the same instant or at point E on curve B. As the current goes through the other half cycle and flows in the opposite direction, a voltage with the opposite polarity will be produced across the resistor. When the current reaches its maximum value at point G on curve A, the voltage will reach its maximum value with the opposite polarity at point G on curve B. As the last quarter of the cycle is completed and the current curve A drops from G to I, the voltage curve B will also drop from G to I.

Notice that throughout the entire cycle the voltage across the resistor was exactly in step with the current flowing through it. We say that the current and voltage are in phase.

Now let us consider what happens across the coil. We already mentioned that in curve A at point A, the current is changing at its maximum rate. Current is leaving terminal 1 of the generator and flowing around the circuit back to terminal 2. This means the current will try to flow through the coil from terminal 1 to terminal 2. The instant the current tries to build up through the coil a voltage will be induced in that coil which will oppose the change in current through it. The amplitude of the voltage will depend on the rate at which the current is trying to change. Since at point A on curve A, the change in current is at its maximum value, the maximum voltage will be built up across L. The polarity of the voltage will be such that it will oppose the current flowing in the circuit. This means that we could get the same effect by putting a battery in the circuit that would prevent the current from flowing through the coil. In order to do this we would have to connect the battery so that terminal 1 was negative and terminal 2 positive; this would oppose the current trying to flow around the circuit.

In Fig. 20C we have represented the voltage at the beginning of the current cycle as A. Notice that the voltage is at its maximum value because the rate of change of current as shown on curve A is at its maximum value at point A.

As the current in the circuit increases from A to B, the rate at which it is changing decreases. This means that the voltage induced in the coil will decrease until the current cycle has reached point B on curve A, and the voltage will have reached point B on curve C. Finally, when the current shown on curve A reaches point C where its rate of change has dropped to 0, the voltage will have dropped to point C on curve C and since the rate of change of current is 0 the voltage induced in the coil will be 0.

At point C on the current curve, the current begins to decrease in value. The rate at which it decreases begins to increase from point C to point D and reaches its maximum value at point E. Since the current is decreasing, the voltage induced in the coil will have the opposite polarity because it will try to prevent this decrease. It will reach maximum value at point E where the rate of current change is at a maximum. This is shown by the voltage waveform between points C, D and E on curve C. From point E to point G the current is still changing in the same direction, but its rate of change is decreasing from E to F until finally when it reaches G, its rate of change is 0. The voltage induced in the coil is represented by the portion of the curve between point E, F and G on curve C. Notice that once again when the current wave has reached point G where its rate of change is 0, the voltage waveform will also be at G. Notice that it is the rate of change of current which controls the voltage induced in the coil, not the actual value of the current flowing in the coil.

From point G to I on the current waveform shown in Fig. 20A, the current begins to change and the rate of change increases until it reaches its maximum value at point I. The voltage waveform is shown at C and the amplitude of the voltage increases from point G where the rate of current change is 0 to the maximum value at point I where the rate of current change is at a maximum.

From examining curves A and C, we can see that the changes in current and voltage across the coil do not occur at the same instant. As a matter of fact, since the current waveform from A to I represents one complete cycle, from A to E and from E to I represents a half cycle. We also refer to this as 180°. (There are 360° in a circle and half a circle is 180°.) From point A to C is one quarter cycle as is from C to E, from E to G and from G to I. Notice that the voltage waveform is identical to the current waveform except that it is one quarter of a cycle ahead of the current waveform. In other words, as the current starts at point A to build up to a maximum value at point C, the voltage is already at its maximum value at point A and starts to drop to its minimum value at point C. During the next quarter cycle when the current drops from its maximum value with one polarity at point C to point E, the voltage is ahead of it by one quarter of a cycle and goes from 0 to its maximum value with the opposite polarity at point E. We say that the current and voltage are out of phase. We refer to this as one quarter of a cycle or 90° phase difference. Since the voltage is ahead of the current we say that the voltage leads the current by 90° or one quarter cycle.

Summarizing what we have seen from Fig. 20, we notice that in the case of a resistance the voltage and current are in phase, but in the case of a coil the voltage leads the current by 90°. In any pure inductance, the voltage will always lead the current by 90° - this is an extremely important point; be sure that you remember it. We can also say that the current lags the voltage by 90° - this is the same thing as saying the voltage leads the current by 90°; the voltage is ahead of the current, therefore, the current must be behind or lagging the voltage.

Now what about the generator voltage - so far we have considered only the generator current. What is the phase relationship between the generator voltage and the generator current?

To simplify our problem let us assume that R has a resistance of 1000 ohms and that L has an inductive reactance of 1000 ohms. Therefore, any current flowing through R and through L will produce equal voltages across them. The voltage across the resistor will be IR and the voltage across the coil will be IX_L.

The waveforms in Fig. 20 tell us that the voltage across the coil is not in phase with the voltage across the resistance. This means that the two do not have their maximum values of voltage at the same instant nor do they have their minimum values at the same instant. Therefore, since the voltages are ac voltages and are not occurring at the same time, we can't simply add them together to find the total voltage across the two. For example, when the voltage across the resistor is at its maximum value as shown at point C and G in Fig. 20B, the voltage across the coil will be at 0. Similarly, when the voltage across the coil is at its maximum value as shown at point A, E and I in Fig. 20C, the voltage across the resistance is 0.

The relationship between the current, resistor voltage, coil voltage and generator voltages can be shown by means of a diagram called a vector diagram. The first step in drawing a vector diagram is to draw a vector to represent the current. We usually draw a horizontal line with an arrow on it and label it I to represent the current.

We know that the voltage across the resistor is in phase with the current and therefore we draw a vector E_R to represent the resistor voltage and this vector will fall on top of the current vector as shown in Fig. 21A. We can select any arbitrary length for this vector since we do not know the generator voltage or the current flowing in the circuit.

We know that the voltage across the coil will lead the current by 90°. Therefore, we draw another vector E_L which is rotated 90° counterclockwise from the current vector as shown in Fig. 21A. Since the value of the resistance of R is equal to the inductive reactance of L, the voltage across the coil will be equal to the voltage across the resistor, therefore, we draw E_L the same length as E_R. The diagram shown in Fig. 21A represents the voltage across the resistor and the voltage across the coil.

To find the generator voltage we complete the vector diagram as shown in Fig. 21B. In this diagram we have drawn a dotted line from the end of the vector E_R parallel to vector E_L. We have drawn another dotted line from the end of vector E_L parallel to vector E_R. The point at which the two vectors intersect gives the value and phase relationship between E_G and I. Since E_L and E_R are equal, the angle between E_G and I will be a 45° angle, and the length of E_G will be 1.4 times the length of either E_L or E_R. This means that the generator voltage will be 1.4 times the voltage across either the coil or the voltage across the resistor.

Perhaps you noticed an apparent contradiction between what we have learned about the voltages in the circuit shown in Fig. 19 and Kirchhoff's Voltage Law. You will remember that Kirchhoff's Voltage Law stated that the sum of the voltage drop in a closed circuit is equal to the source voltage. If this is true for ac circuits then the voltage drop across R plus the voltage drop across L must be equal to the generator voltage. Yet in Fig. 21 we found that the generator voltage was only 1.4 times either E_R or E_L. At first glance you might think it should be twice E_R or E_L. However, remember that E_R and E_L are not in phase. This means that if you add these voltages at any

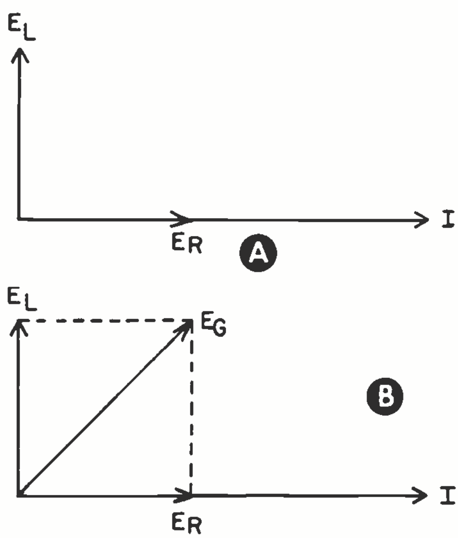


Fig. 21. Vector diagrams showing relationship between current, generator voltage, coil voltage and resistor voltage of the circuit shown in Fig. 19.

instant their sum will be equal to the generator voltage.

IMPEDANCE

In Fig. 19, we considered the coil as a pure inductance. We treated the resistance of the wire used to wind the coil as 0; the only opposition the coil offered was inductive reactance. Of course the wire will have resistance as well as certain other resistive effects to the ac current. The ac resistance of the coil is the sum of the resistance of the wire plus these other losses which increase the resistance. The impedance of the coil is the total opposition to current flow. It is made up of the opposition due to the inductive reactance of the coil and the opposition due to the ac resistance of the coil.

The practical way of studying how a coil behaves in an ac circuit is to consider the coil as being made up of a pure inductance with a resistor in series with it. This is essentially the type of circuit used in Fig. 19. Here you see that in the circuit the current lags behind the generator voltage. Where the resistance was equal to the inductive reactance, the phase difference was 45°. In actual practice, the resistance will usually be much smaller than the inductive reactance so that the phase difference will approach 90°. The higher the ratio of inductive reactance to resistance, the closer the phase difference will approach 90°.

In studying a complete circuit in which there is a coil and a resistance in the circuit, you can simply lump together the resistance of the resistor and the coil resistance and consider this the resistance in the circuit and then treat the inductance of the coil separately. However, if you are interested in finding the voltage across the coil, then you have to keep the resistance of the coil separated and use it with the inductive reactance of the coil since both will have an effect insofar as developing voltage across the coil is concerned.

FINDING THE CURRENT IN AN AC CIRCUIT

If you want to find the current flowing in this type of circuit when an ac voltage is applied, you must find the impedance of the circuit. The impedance is the total opposition to the ac current flow in the circuit.

There are several ways of finding the total flow of ac current in the circuit. We have already briefly started to introduce one in Fig. 21. We will go through this procedure in detail now and then show you another method. You may use whichever way is easier for you.

As an example, let us find the current flowing in the circuit shown in Fig. 22. Here we have a coil with an inductance of 2 henrys. This coil is connected in series with a 1000-ohm resistor. The two are connected across a 60-cycle generator having an output voltage of 500 volts. The resistance of the coil is so small, that compared to the 1000-ohms in the circuit, it is insignificant so we can ignore it. The problem is to find the current that will flow in the circuit.

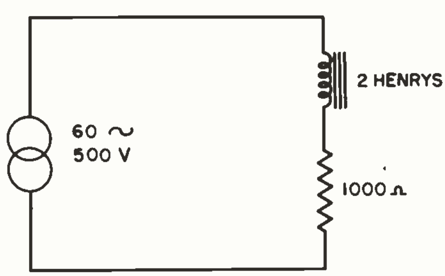


Fig. 22. There are two ways to find the current in the circuit shown above: by using vectors or by using a mathematical solution.

Vector Solution.

First, we must find the inductive reactance of the coil. To do this we use the formula:

$$X_L = 6.28 \times f \times L$$

and substituting 60 for f and 2 for L we get:

$$X_L = 6.28 \times 60 \times 2$$

which equals 753.6 ohms. Since this is a practical problem, we simply call it 750 ohms.

Now we know that the inductive reactance of the coil is 750 ohms and the resistance in the circuit is 1000 ohms. You might at first think that we can obtain the total opposition to the ac current flow simply by adding these two together. However, this is not true--you cannot simply add inductive reactance and resistance. Let us see why. We know that when voltage is applied to an inductance, the current that flows will be out of phase with the applied voltage. When voltage is applied to the resistance, the current that flows will be in phase with the applied voltage. The inductance and the resistance have different effects on current.

Adding the effect of the two together can be done by means of vectors.

You have already seen how vectors

can be used to indicate phase differences in quantities having the same frequency. Now we will see how they can be used to add similar quantities having the same frequency but a difference in phase.

As before, the angle between the vectors represents the phase difference between the quantities. The arrows are all drawn to the same scale, so that the length of the arrows indicates the amplitudes of the quantities to be added.

For example, suppose we wanted to show the relationship between two 60-cycle ac voltages A and B. A is 30 volts, and B is 40 volts, and they are 90° out of phase with each other.

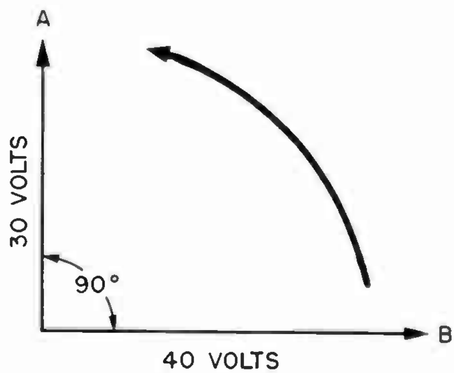


Fig. 23. In this diagram, the lengths of the arrows show the amount of voltage and the angle between them shows their phase relationship. They are considered to rotate counterclockwise.

A is leading B. Fig. 23 shows how we would draw this, using a scale of 1/2-inch equals 10 volts. We draw B 2 inches long, and we draw A, 1-1/2 inches long. Since A is leading B by 90°, we draw it 90° counterclockwise from B.

Now, suppose we wanted to find the sum of these voltages. We could not simply add 30 and 40, because

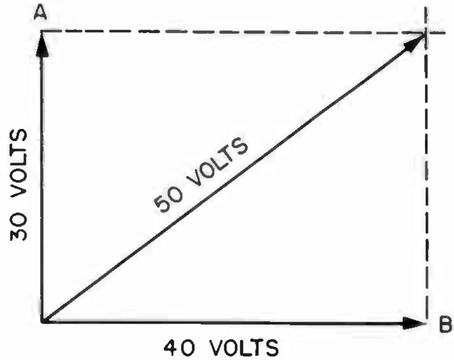


Fig. 24. How to find the vector sum of two ac voltages differing in phase.

of the difference in phase. This is where the vector diagrams will help us. We can add these two voltages, taking into account the phase difference as shown in Fig. 24. We say we are finding the "vector sum" of the two.

To do this, we complete a rectangle by drawing lines parallel with the two vectors. Then we draw in a diagonal to the point where the two lines intersect. This diagonal represents the vector sum of voltages A and B. When we measure it, we see it is 2-1/2 inches long. Since we used the scale of 1/2 inch to 10 volts, we see that the vector sum is 50 volts.

Now let us see how we can apply this principle to find the total opposition or impedance in the circuit we have been studying. As we have already mentioned, when considering phase in a circuit, the phase of the current is always used as a reference. The current vector is drawn horizontally and pointing to the right, and voltage vectors are drawn in the positions corresponding to their phase relationship to the current. So the first thing we do is to draw an arrow to represent the current, as

Fig. 25. The current vector is drawn horizontally and used as a reference point for the other vectors.

shown in Fig. 25. We do not know what the current is, so its length does not matter, but we do know it should be drawn horizontally and pointing to the right. The next step in our procedure depends upon an important fact--that the voltage across the resistor will be in direct proportion to its resistance. We also know that the voltage across the resistor will be in phase with the current, and the voltage across the coil will be 90° ahead of the current. Therefore, we can draw two vectors, one on top of the current vector to represent the voltage across the resistor, and one 90° ahead of (counterclockwise from) the current vector to represent the voltage across the coil. We do not know what these voltages are, but since the voltage is in direct proportion to the resistance and reactance, we can draw the arrows using a scale that is in proportion to the ohmic values of the resistance and reactance and label the vectors R and X_L.

Fig. 26. The voltage across a resistor is in phase with the current, so the vector for the resistor voltage is drawn on top of the current vector.

First we draw a vector to represent the voltage across the resistance. If we use the scale of an inch to 500 ohms, the vector representing the voltage across the resistor will be 2 inches long. Since the current flowing through the resistor will be in phase with the voltage, we draw the resistance voltage vector and mark it R as shown in Fig. 26. Here you should notice that it is drawn right on top of the current vector. The current vector is not drawn to scale, but the resistance voltage vector is drawn 2 inches long.

Next, we draw the vector for the voltage drop across the coil. Since we have a reactance of 750 ohms, this vector should be 1-1/2 inches long. Since the coil voltage is 90 degrees out of phase with the resistor voltage, this vector is drawn as shown in Fig. 27 and labeled X_L. Now we have a vector diagram that represents the voltage across the resistance and the reactance in the circuit shown in Fig. 22. To get the impedance, we draw dotted lines as shown in Fig. 28 to complete the rectangle. The vector representing the voltage across the impedance is drawn in as shown in Fig. 28 to the point where these lines meet, and the impedance is obtained by measuring the length of this vector. On the diagram we have drawn, the impedance voltage vector is 2-1/2 inches long. Since we have used the scale of 500 ohms to the inch, the impedance in the circuit must be 2-1/2 times this value, or 1250 ohms. This is the total impedance or opposition to current flow in the circuit. Now that we have this figure, we can quickly determine the current that will flow.

To find the current we use Ohm's Law for coils. The current is equal to the voltage divided by the impedance. The letter Z is usually used to represent impedance. This can be expressed

$$I = \frac{E}{Z}$$

E=IZ

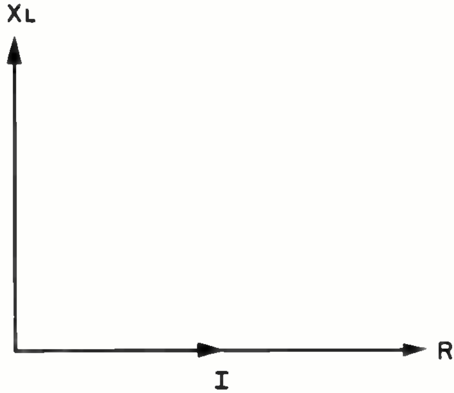


Fig. 27. The vector for the voltage across the inductance is drawn in at right angles to the resistor voltage vector.

Substituting 500 volts for E and 1250 ohms for Z we get:

$$I = \frac{500}{1250} = .4 \text{ amp}$$

Mathematical Solution.

Another method of solving for the impedance in an ac circuit is by means of the formula:

$$Z = \sqrt{R^2 + X_L^2}$$

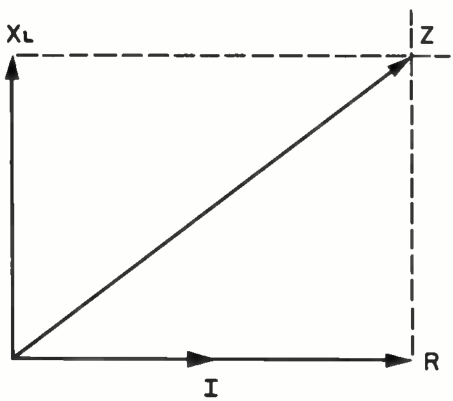


Fig. 28. The vector representing the voltage across the impedance is the vector sum of the other two.

The mathematical sign $\sqrt{\ }$ means to find the square root. Therefore, you square the resistance and the reactance, add the two together, and then take the square root of the sum. Once you have the impedance, proceed as before to get the current. Again,this is not the type of problem that the technician will have to solve, but it is important to remember the general method of obtaining the impedance in a circuit of this type. If you know how to do square root problems, the mathematical solution is the simpler; if you don't, the graphical solution is the one to use. It is particularly important to realize that you cannot obtain the impedance simply by adding the resistance and the reactance together. The impedance in a circuit will always be somewhat less than the sum of the two because of the difference in phase.

KIRCHHOFF'S VOLTAGE LAW

You will remember that Kirchhoff's voltage law stated that the sum of the voltage drops in a complete circuit is equal to the source voltage. Now let's see how this applies to an ac circuit consisting of inductance and resistance.

Using the same example as before, we have already calculated the reactance of the coil at 750 ohms and we know the resistance of the resistor is 1000 ohms. We can use Ohm's Law in the form:

$$E = I \times R$$

to find the voltage drop across the resistor. In the case of the coil, the voltage drop is:

$$E = I \times X_L$$

To find the voltage drop across the coil, we simply multiply 750 by the current, which we have already determined as .4 amp. 750 × .4 = 300. Therefore the voltage across the coil is 300 volts. The voltage across the resistor is 1000 ×.4 = 400 volts. Now look at Fig. 29 where we have indicated the voltages. We have 300 volts across the coil and 400 volts across the resistor, but our source voltage is only 500 volts. These are the readings we would actually obtain if we had meters connected as shown!

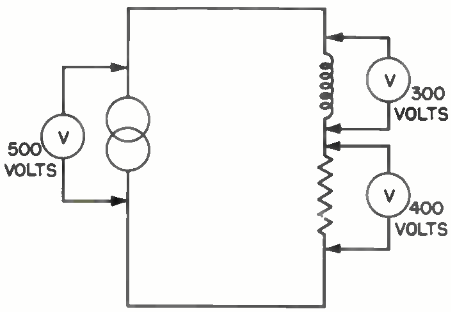


Fig. 29. With a source voltage of 500 volts, we have 300 volts across the coil, and 400 volts across the resistor, but they are not in phase.

This may appear to be a contradiction of Kirchhoff's Voltage Law, but actually it is not. You must remember that the voltage across the resistor will be in phase with the current, but the voltage across the coil will not be in phase with the current flowing through it. The voltages that we have just determined are effective voltages. At any given instant the sum of the voltage across the coil plus the voltage across the resistor will be equal to the voltage across the generator. However, the effective voltage across the coil and the resistor if they are simply added together would give us more than 500 volts.

To add these two voltages we must again resort to vectors. The vector addition of these two voltages using a scale of 200 volts equals 1 inch is shown in Fig. 30. Notice that the vector representing the voltage across the resistor is drawn 2 inches long, and the one representing the voltage across the coil is drawn 1-1/2 inches long. When we complete the vector diagram to find the sum, as before, we obtain a vector which is equal to the generator voltage of 500 volts.

The mathematical solution that we used before can also be used to obtain the source voltage. If we let the symbol E_G equal the source voltage we have:

$$E_G = \sqrt{E_R^2 + E_L^2}$$

If we substitute 400 for E_R, E_R^2 equals 160,000. Similarly E_L squared equals 90,000. Adding the two together we get 250,000; the square root of 250,000 is 500, so as before we find that $E_G = 500$ volts.

Importance of Phase.

A general knowledge of phase and vector diagrams helps you to understand the action of coils and capacitors in ac circuits and will make you a better-than-average technician. You will understand why you do certain things when making adjustments or repairs instead of just blindly following instructions. It is the men who know the "How" and the "Why" who command the highest salaries in the modern world of electronics.

Later in your course, you will learn that the opposition of a coil, which we have called inductive reactance, can be balanced or cancelled by the opposition of a capaci-

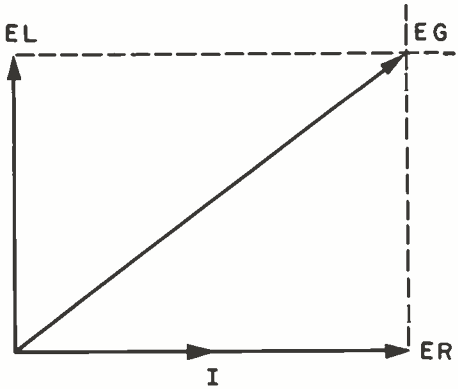


Fig. 30. Finding the vector sum of ER and EL.

tor, which is called capacitive reactance, because the two are of opposite phase. You will see the importance of phase in many other practical examples. Phase, however, is not a subject you can grasp in one lesson; you will understand it better and better with each succeeding lesson.

Q OF A COIL

We mentioned that since a coil is wound of wire and wire has resistance, there is no such thing as a perfect inductance. All coils have both inductance and resistance.

As you might expect, when manufacturers make a coil, they usually try to keep the resistance as low as possible. Generally, the lower the resistance is in proportion to the inductive reactance of a coil, the better the coil. The relationship between inductive reactance and resistance is called the Q of the coil. This is represented by the formula:

$$Q = \frac{X_L}{R}$$

A high-Q coil is a coil in which the value of the inductive reactance is much higher than that of the resistance. Coils with a Q of 100 or more are quite common.

Since the reactance of a coil increases with frequency, you might expect the Q to increase with frequency. This is true up to a certain point, but R is the ac resistance of a coil and it increases with frequency also. As long as X_L increases with frequency faster than R, the Q of the coil will increase, but if R increases faster than X_L, the Q of the coil will decrease as the frequency increases so the coil cannot be used at high frequencies.

Q is particularly important in tuned circuits when coils are used with capacitors. You will see why this is so later when you study these circuits.

SUMMARY

This section of your lesson is almost too important to try to summarize. However, to help you to review, here are the important things you should understand.

You should have a general understanding of what we mean by phase. When the current in an ac circuit is increasing exactly in step with the voltage, and reaches the maximum value at the same time as the voltage reaches the maximum value and reaches its minimum value at the same time as the voltage reaches its minimum value, we say that the current and voltage are in phase. In a circuit consisting of a pure inductance the current will lag the voltage by 90 degrees. This means that it is one-quarter of a cycle behind the voltage.

Impedance is the vector sum of resistance and reactance. The impedance in a circuit will be greater than the resistance or the reactance alone. Impedance cannot be determined simply by adding the resistance and the reactance.

The voltage across a component in an ac circuit can be found by Ohm's Law. The sum of the individual voltage drops in an ac circuit is equal to the source voltage, providing we add these voltages vectorially. We cannot add them by means of simple arithmetic and expect their sum to be equal to the source voltage. If we could measure the source voltage and the voltage across each of the parts in the circuit at any instant, we would find that the sum of the voltage drops at that instant would be equal to the source voltage.

SELF-TEST QUESTIONS

(ak) What do we mean when we say that the voltage and current in a circuit are in phase?

(al) What is the phase relationship between the voltage and current across a resistor?

(am) What is the phase relationship between the voltage and current across a coil?

(an) At what point in an ac cycle is the current changing at its maximum rate?

(ao) What is meant by impedance?

(ap) An ac generator with an output voltage of 250 volts is connected across a 1.5-henry coil and a 900-ohm resistor in series. The frequency of the generator is 100 cycles. Find the current flowing in the circuit.

(aq) In the preceding problem, find the voltage across the coil and the voltage across the resistor.

LOOKING AHEAD

You have now finished the study of the basic facts of resistors and coils. When you complete a similar study of capacitors in the next lesson, you will have a basic knowledge of these three important parts. In a later lesson you will learn more about how these three parts work together and what effect they have on ac signals. Remember that the ac supplied by the power company, audio signals, and radio frequency signals are all ac signals differing only in frequency and in some cases in wave shape. The important facts you learned about ac and coils in this lesson apply to all ac signals regardless of their frequency.

Most students are anxious to go ahead as quickly as possible with their course, particularly in the early lessons. However, do not be so anxious to go ahead with later lessons that you leave the earlier lessons without completely understanding them. The information given in these early lessons is basic and is information that you will use over and over again in more advanced lessons. If you do not understand how basic parts such as resistors, coils, and capacitors affect circuit performance, you will not be able to understand some of the later lessons.

ANSWERS TO SELF-TEST QUESTIONS

(a) Magnetic lines of force form complete loops.

(b) Near the coil.

(c) An air-core coil. The cardboard form merely supports the turns of the coil; it has no appreciable effect on the operation of the coil.

(d) Laminations are thin strips of iron or steel used to produce an iron core for a choke or transformer.

(e) An iron core provides a better path for the magnetic lines of force. We say it has a lower reluctance.

(f) Magnetomotive force is the force that sends magnetic flux around a magnetic circuit.

(g) The ampere-turn.

(h) 50 ampere - turns.

(i) Reluctance is the opposition to flux in a magnetic circuit. It is the equivalent of resistance in an electrical circuit.

(j) Yes. The reluctance in the magnetic circuit of an iron-core coil is much lower than the reluctance of the magnetic circuit in an air-core coil.

(k) They all have the same reluctance. Non-magnetic materials have the same reluctance as air.

(l) The permeability of a material indicates the ability of the material to pass the magnetic lines of force. The higher the permeability of the material, the less reluctance it will offer to magnetic lines of force.

(m) In a magnetic circuit the flux is equal to the magnetomotive force divided by the reluctance.

(n) A flux linkage is a magnetic line of flux cutting through a single turn of a coil. If the magnetic line of flux cuts through two turns then we have two flux linkages, and if it cuts through five turns then we have five flux linkages.

(o) Twelve. A magnetic line of flux cutting through a single coil produces one flux linkage. Therefore three lines cutting through four turns produces $3 \times 4 = 12$ flux linkages.

(p) Zero. If the number of flux linkages cutting a coil does not change, there will be no voltage induced in the coil.

(q) The induced voltage will produce a current which will in itself produce a magnetic field which will oppose any change in flux linkages.

(r) It will aid the original lines of force. If the field is reduced, the induced voltage produced in the coil will cause a current to flow in such a direction that the magnetic field produced will tend to prevent the number of flux linkages from decreasing. In order to do this it must aid the original field.

(s) (1) cutting lines of force, (2) changing the reluctance (3) changing the coil current.

(t) The variable reluctance phono pickup. In this type of pickup the needle moves between two coils. As the needle follows the record groove it moves from side to side decreasing the air gap on one side so that the reluctance on that side decreases and the flux increases.

(u) The transformer. Varying ac applied to the primary of the transformer causes a varying current to flow. This current in turn causes a varying flux; the varying flux cutting the secondary induces a voltage in the secondary of the transformer.

(v) Inductance.

(w) A self-induced voltage is a voltage induced in a coil that is caused by a change in the current flowing through the coil. The changing current causes a change in magnetic flux. The change in magnetic flux produces a self-induced voltage which tends to produce a current which in turn will produce a magnetic field opposing the original change.

(x) It will oppose the applied voltage. The self-induced voltage will try to keep the current constant so that the flux will remain constant. To do this it must oppose the applied voltage.

(y) The henry is the unit of inductance.

(z) If the voltage induced in a coil is 1 volt when the strength of the current flowing through the coil changes at a rate of 1 ampere per second, the coil has an inductance of 1 henry.

(aa) (1) the number of turns on the coil, (2) the diameter of the coil, (3) the permeability of the core material.

(ab) The inductive reactance of a coil is the opposition the coil offers to the flow of ac through it due to the inductances of the coil.

(ac) The inductive reactance of a coil is measured in ohms.

(ad) 3768 ohms. To find the inductive reactance of a coil you use the formula:

$$X_L = 6.28 \times f \times L$$

and substituting 60 for f and 10 for L we have:

$$X_L = 6.28 \times 60 \times 10$$
$$X_L = 3768 \text{ ohms.}$$

(ae) When two coils are placed near each other so that the flux from one coil cuts through turns of the other coil, any change in the flux from one coil will induce a voltage in the other coil. We call this coupling between the two coils mutual inductance.

(af) 14 henrys. When two coils that

are not mutually coupled together are connected in series, the total inductance is simply the sum of the two ininductances.

(ag) 11 henrys. To find the total inductance of the two coils we use the formula:

$$L_T = L1 + L2 + 2M$$

and substituting 4 henrys for L1 and 3 henrys for L2 and 2 henrys for M we get:

$$L_T = 4 + 3 + (2 \times 2)$$
$$L_T = 11 \text{ henrys.}$$

(ah) 9 henrys. To find the inductance of the two coils we use the formula:

$$L_T = L1 + L2 - 2M$$

and substituting 8 henrys for L1, 7 henrys for L2 and 3 henrys for M we get:

$$L_T = 8 + 7 - (2 \times 3)$$
$$L_T = 9 \text{ henrys.}$$

(ai) 2200 millihenrys. To convert henrys to millihenrys you just multiply by 1000 which is the same as moving the decimal point three places to the right. $2.2 \times 1000 = 2200$.

(aj) .0001 henry. To convert microhenrys to henrys you divide by 1,000,000 or move the decimal point six places to the left. Moving the decimal point six places to the left requires that we add three zeros to the left of 1 and then simply move the decimal point.

(ak) When we say that the voltage and current are in phase we mean that any change in voltage produces a corresponding change in current. In other words, an increase in voltage causes an instant increase in current or a decrease in voltage causes an instant decrease in current.

(al) The voltage and current across a resistor are in phase.

(am) The voltage across a coil will lead the current by 90°. Another way of expressing the same thing is to say that the current lags the voltage by 90°.

(an) The current is changing at its maximum rate when the current wave is going through 0.

(ao) Impedance is the total opposition to current flow. It is made up of the reactive opposition and the resistive opposition to current flow.

(ap) $X_L = 6.28 \times 100 \times 1.5$
$= 942.0$ ohms

$$Z = \sqrt{R^2 + X_L^2}$$
$$= \sqrt{900^2 + 942^2}$$
$$= \sqrt{810,000 + 887,364}$$
$$= \sqrt{1,697,364}$$

$= 1303$ ohms

$$I = \frac{E}{Z} = \frac{250}{1303} = .19 \text{ amps}$$

(aq) $E_R = IR$

$= .19 \times 900 = 171$ volts

$E_L = IX_L$

$= .19 \times 942 = 179$ volts

Lesson Questions

Be sure to number your Answer Sheet B106.

Place your Student Number on every Answer Sheet.

1. If the reluctance of a magnetic circuit is increased, what will happen to the flux?

2. If the number of flux linkages cutting a coil decreases, a voltage will be induced in the coil that will cause a current to flow, which will produce magnetic flux that will tend to prevent the original flux from decreasing. Is this statement true or false?

3. Explain what inductive reactance is.

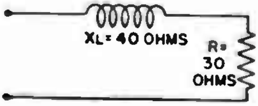
4. What is the inductive reactance of a 1-henry coil at a frequency of 100 cycles?

5. If two 15-henry coils have a mutual inductance of 5 henrys, what is the total inductance when they are connected in series if the flux of one coil aids the flux of the other?

6. Explain what is meant when we say "the voltage and current are out of phase."

7. What do we mean by the impedance of a circuit?

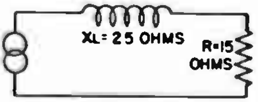
8. What is the impedance of the circuit shown?



9. If the frequency of a voltage source connected to a circuit consisting of a coil and resistor in series is increased, will the current flowing in the circuit increase, decrease, or remain the same?

10. If the current flowing in the circuit shown is 1 amp, find:

    (1) the voltage across the coil
    (2) the voltage across the resistor.



SINCERE APPRECIATION PAYS

Have you ever watched a dog respond to a friendly pat as a reward for obedience? Have you noticed how a child glows with joy when praised for good behavior? Have you ever felt your own brain cells respond with increased effort when you praise them by saying, "That's a fine piece of work, even if I did do it myself!"

Yes, everyone responds to sincere and merited praise. It is a tonic to both giver and receiver. It brings greater praise and appreciation back to you. It costs nothing more than a smile and a few sincere words, but it can truly achieve miracles in happiness and success, and put real money in your pocket.

Time spent in figuring how to give sincere and deserved praise is well worth while. Let people know that you appreciate their fine work, and watch the breaks come your way.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-206

ACHIEVEMENT THROUGH ELECTRONICS

HOW CAPACITORS ARE USED

B107

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

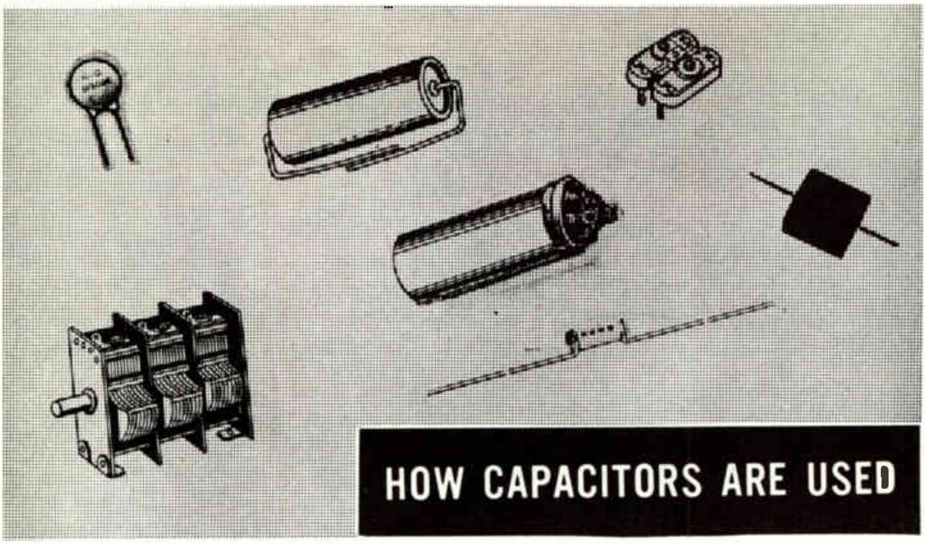
HOW CAPACITORS ARE USED

B107

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

PI12M375 1975 EDITION Litho in U.S.A.



HOW CAPACITORS ARE USED

Of the three major electronic circuit parts, resistors, coils and capacitors, it would be impossible to pick the one that is the most essential. All three parts are extremely important. In many circuits all three are used together; in some circuits two of the three are used together. When used in combination, these parts are able to perform jobs that one cannot do alone.

In the preceding two lessons, you studied resistors and coils. In this lesson you will study capacitors in detail. After you have completed this lesson, you should have a good understanding of how these three basic parts work. Later, you will see how they are used together.

In many respects a capacitor is the opposite of a coil. You will remember that in an ac circuit with a coil in it, the current flowing in the circuit will lag or follow the applied voltage by 90 degrees. If instead of a coil we put a capacitor in the circuit, the current would lead the voltage by 90 degrees; in other words we would have just exactly the opposite effect. You will see more of this later in this lesson and also find out exactly why it is that in a capacitive circuit the current leads the voltage. For the present, let's start this lesson by learning a little about what a capacitor is and some of the fundamentals of how it works.

WHAT IS A CAPACITOR?

In its simplest form, a capacitor, or condenser, as many old-timers call it, consists of nothing more than two pieces of metal separated either by air or by some other non-conducting material placed between them. The material between the two plates, whether it is air, a liquid, or a solid, is called the dielectric. If there is nothing between the plates but air, we say the capacitor has an air dielectric.

The electrical size of a capacitor is called its capacity. A large ca-

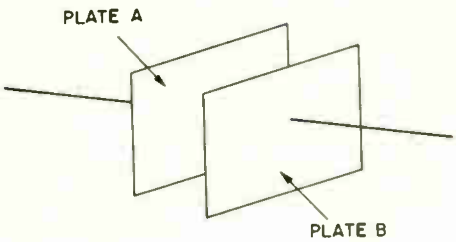


Fig. 1. A simple capacitor is nothing but two pieces of metal separated from each other.

pacitor has a large capacity. There are a number of things that affect the capacity of a capacitor, which you will study in a little while.

HOW A CAPACITOR WORKS

A simple capacitor made of two metal plates with an air dielectric between them is shown in Fig. 1. To see how a capacitor works, let's see what will happen when we connect this capacitor to a battery as shown in Fig. 2.

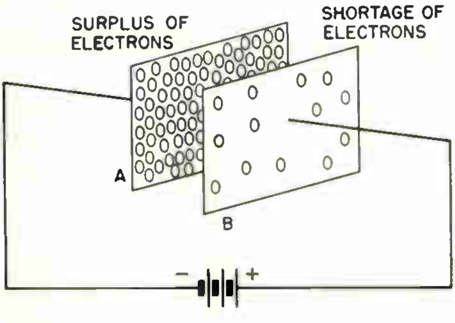


Fig. 2. When a capacitor is connected to a battery, a surplus of electrons is accumulated on one plate. This forces electrons off the other plate, leaving it with a positive charge.

When the plates of the capacitor are connected to the battery, electrons flow from the negative terminal of the battery into the plate of the capacitor connected to the negative terminal. There will be a surplus of electrons built up on this plate of the capacitor.

You know that one of the characteristics of an electron is that it repels other electrons. Remember the rule of charges, like charges repel. Therefore, the surplus electrons on the one plate of the capacitor will repel electrons from the other plate back to the positive terminal of the battery. At the same time, the positive terminal of the battery attracts electrons and pulls them from the plate connected to it, leaving a shortage of electrons on this plate, giving it a positive charge. This positive charge will attract electrons from the negative terminal of the battery to the plate connected to it. Thus, there will be a surplus of electrons on one plate and a shortage of electrons on the other. The electron flow will continue until plate A is just as negative as the negative battery terminal, and plate B is just as positive as the positive battery terminal. When this condition exists we say the capacitor is charged.

If we suddenly disconnect the capacitor from the battery, the condition of unbalance that has been set up on the capacitor plates will remain. We will have a surplus of electrons on one plate and a shortage of electrons on the other. Thus, we have electricity stored in the capacitor.

You will remember that one of the characteristics of a charged object is that it tries to give up its charge in order to become neutral as quickly as possible. Therefore, if we connect a wire from one plate of a capacitor to the other, the electrons will flow from the side having a surplus of electrons over to the side having a shortage of electrons until the number of electrons on the two plates is balanced, and there is no longer a charge on them.

This is a very brief explanation of how a capacitor works, how it is charged, and how it can store electricity. We will look into this more thoroughly in the next section of this lesson, but this is enough to give you a general idea of how a capacitor works. Keep in mind that a capacitor can store electricity. Before touching the leads of a large capacitor you should short the leads together with a screwdriver or similar object to be sure the capacitor is discharged, otherwise you may discharge the capacitor and receive an unpleasant and possibly dangerous shock!

TYPES OF CAPACITORS

Capacitors can be divided into two types, according to what type of material (called the dielectric) separates the plates. One type has an air dielectric, and the other has a solid or liquid dielectric. When we say a capacitor has an air dielectric, we simply mean that there is nothing but air between the plates of the capacitor. When we say a capacitor has a solid or liquid dielectric we mean that some insulating material other than air has been inserted between the plates.

You will see typical examples of all these types of capacitors later; you will learn more about them, what they look like, and where they are used in electronic circuits; but first let us learn more about how they work.

How Capacitors Store Electricity

In considering a capacitor, you may at first wonder how a capacitor can be used in an electronic circuit because there is no complete circuit through the capacitor. In the sketch of the simple capacitor shown in Fig. 1 you can see that the two plates of the capacitor do not touch each other. There is a space between the two plates so that the electrons on one plate cannot normally flow from one plate to the other.

When we connected a battery to a capacitor we saw that the plates of the capacitor became charged, one plate picking up a surplus of electrons and the other losing electrons so that it had a shortage.

The usefulness of a capacitor depends upon its ability to store electricity or to hold a charge. Let's learn a little more about how a capacitor is charged, so we can better understand some of its more important uses.

CHARGING A CAPACITOR

A capacitor cannot be charged instantly. It takes time for the charge to build up after the electrons start to flow from the negative terminal of the battery into one plate of the capacitor and from the other plate to the positive terminal of the battery. The length of time that it takes depends upon two things, the size of the capacitor and the amount of resistance in the circuit.

You might think that there was no resistance in the circuit we have shown in Fig. 2. However, this is not the case. There is resistance in the leads used to connect the capacitor to the battery, and in addition there is the internal resistance of the battery itself. These two resistances will limit the rate at which the capacitor can charge.

Because it does take some time to charge a capacitor, there will be a current flowing in the circuit shown in Fig. 2 when the capacitor is first connected to the battery. This current will flow as long as the battery is charging the capacitor. The longer it takes the battery to charge the capacitor the longer there will be a current flowing in the circuit. Therefore, you can see that even though the electrons cannot cross from one plate of the capacitor to the other plate there is a current flow in the circuit, at least for the short time it takes to charge the capacitor.

A question that sometimes comes up when considering a charged capacitor is whether or not the capacitor has any more electrons than it has in the discharged state. The answer to this question is no--the capacitor will have the same number of electrons whether it is charged or discharged. The only difference is that when a capacitor is discharged there is no charge on any of the atoms making up the metal on either plate. In other words, each atom has enough electrons to exactly neutralize the charge in its nucleus. However, when the capacitor is charged, some of the electrons are moved off one plate so there is a shortage of electrons on that plate, and the same number of extra electrons are forced onto the other plate, so there is a surplus of electrons on it. Thus the total number of electrons in the material making up the capacitor does not change.

The Amount of Charge.

The charge on a capacitor depends upon the battery voltage used to move electrons onto one plate and away from the other. A battery with a higher voltage can exert more force on the atoms making up the capacitor plates and thus move more electrons than a battery with a lower voltage could. However, there are other things that affect the charge we can store in a capacitor. The electrical size of the capacitor is just as important as the charging voltage. The electrical size of the capacitor is called the capacity of the capacitor. Now let's see what we mean by capacity.

CAPACITY

The term capacity is used to describe the electrical size of a capacitor. It is used in the same way as inductance is used to describe the electrical size of a coil, and resistance is used to indicate the electrical size of a resistor.

Just as the henry is the unit of inductance and the ohm is the unit of resistance, the "farad" (pronounced FAIR-ad) is the unit of capacity. It was named after the scientist Michael Faraday, who did a great deal of the early work with capacitors. The capacity of a capacitor is a measure of its ability to store electricity. A capacitor with a high capacity can store more electrons than a capacitor with a lower capacity. Thus the capacity of a capacitor indicates its electrical size to the technician just as the resistance of a resistor indicates the electrical size of the resistor.

We can express the capacity of a capacitor in terms of charge and voltage. The capacity of a capacitor is equal to the charge it will take divided by the voltage used to put that charge on the capacitor. The amount of charge is expressed in units called "coulombs." A coulomb represents a certain quantity of electrons. If a current of 1 ampere flows in a circuit for one second, the number of electrons moving past a given point in the circuit represents one coulomb of electricity. If when we connect a one-volt battery across a capacitor we can store a charge of 1 coulomb in the capacitor, its capacity is 1 farad. If the capacitor would take a charge of 2 coulombs with an applied voltage of 1 volt, the capacity would be two farads. The farad actually represents an extremely large capacity. It is so large in fact that it is never used in electronics. Let us look at the smaller units of capacity that are used.

Units of Capacity.

Since the farad is so large a unit, capacity in electronic circuits is usually expressed in smaller units, which are fractions of a farad. They are:

1. The microfarad, which is equal to one-millionth of a farad. Microfarads are abbreviated in several ways; the most common abbreviations are μf, mf, and mfd.

2. The picofarad, which is equal to 1-millionth of a microfarad. It

is abbreviated pf. In addition to the picofarad you will also run into the micro-microfarad, which is also equal to 1-millionth of a microfarad. The micro-microfarad was used for many years to designate a millionth of a microfarad, but in recent years the picofarad has replaced it. If you should be looking at a diagram of a modern radio or TV receiver the chances are you'll find that the abbreviation pf has been used to indicate picofarads but if you are working on an older set then you will find micro-microfarad used. The abbreviations used for micro-microfarad are $\mu\mu f$, mmf, and mmfd.

Because all seven abbreviations are frequently found in electronics, you should learn them all. The Greek letter μ is the Greek letter "mu" which is pronounced MEW.

You should have no difficulty in remembering that the prefix micro means one millionth, because you have run into this several times previously. The microfarad is simply one millionth of a farad. Actually, you will not be dealing with farads at all because this is such a large unit and for practical purposes you can consider that the unit of capacity in electronics is the microfarad. The smaller unit, which you will have to deal with, is the picofarad, which is a millionth of a microfarad.

Sometimes it is necessary to change from microfarads to picofarads and vice-versa. To change from microfarads to picofarads, you simply multiply by one million or move the decimal point six places to the right. In other words, a capacitor that has a capacity of 5 microfarads has a capacity of 5,000,000 picofarads. You simply add six zeros. A capacitor that has a capacity of .0005 microfarads has a capacity of 500 picofarads. You simply move the decimal point six places to the right.

To convert from picofarads to microfarads you divide by one million, and this can be done by moving the decimal point six places to the left. To do this, you can add zeros to the left. Thus 100 picofarads can be written 000,000,100 picofarads. All the zeros to the left of the 1 have no meaning. To convert this value to microfarads, move the decimal point six places to the left and you get 000.0001 microfarad. The zeros preceding the decimal point have no meaning, so they can be dropped, and you have .0001 microfarad or .0001 mfd.

Since the micro-microfarad is also equal to one millionth of a microfarad - in other words it is equal to a picofarad - if you should run into an older receiver where the values are given in micro-microfarads and you want to convert the value to microfarads you use exactly the same procedure as used in converting picofarads to microfarads.

No doubt as you are reading the preceding section you noticed how long the words microfarad and picofarad are. Technicians have shortened these words. The word microfarad is frequently abbreviated to "mike". Thus if you went to a wholesaler to buy a 2-microfarad capacitor you would probably simply say, "I want a 2-mike capacitor". Technicians have shortened picofarad to simply the abbreviation pf. Thus if you were ordering a 100 picofarad capacitor you would probably order it as a 100-pf capacitor rather than pronouncing the entire word.

Another abbreviation is as follows: instead of saying decimal point, 0, 0, 1 microfarad to identify a .001 mfd capacitor technicians usually say "point double oh one mike". Similarly for .00025 mfd they would say "point triple oh two five mike".

FACTORS AFFECTING CAPACITY

If you have ever looked into tne back of a radio receiver you have probably noticed what we call a variable capacitor. A variable capacitor consists of a set of fixed plates and a set of plates that can be rotated. As the receiver is tuned across the broadcast band the rotating plates will move into position and mesh between the fixed plates. The actual capacitor in the receiver might have been made of two or three separate capacitors all connected together so that the rotating plates all rotate at the same time. A capacitor with two sections is called a 2-gang capacitor and one with three sections is called a 3-gang capacitor. This type of capacitor is called a variable capacitor because its capacity is varied as you rotate the movable plates and mesh them between the stationary plates. When the plates are completely meshed the capacitor has its maximum capacity and when they are rotated so that they are separated as far as possible the capacitor has its minimum capacity.

There are a number of factors that affect the capacity of a capacitor. So you will understand how the capacity of a variable capacitor is changed and also help you to understand the other capacitors you will encounter in electronics work, we will now discuss some of the factors affecting capacity.

1. Area of Plates.

The capacity of a capacitor depends upon the area of the plates. Thus, if the area of each plate of a simple capacitor such as the one shown in Fig. 1 is doubled, the capacity would be doubled.

There are other ways in which the area can be increased in order to increase the capacity. For example, look at the capacitor shown in Fig. 3A. Notice that instead of a simple capacitor made up of two plates as we had in Fig. 1, here we have three plates. Two of the plates, marked A1 and A2, are connected together. If we start off with a capacitor having the two plates A1 and B, and then add the plate A2 to the capacitor, we would double the capacity. You can see why this is so when you consider what happens to the area of the plates when we add plate A2. Let's assume that each plate has an area of one square inch. Thus the area of plate

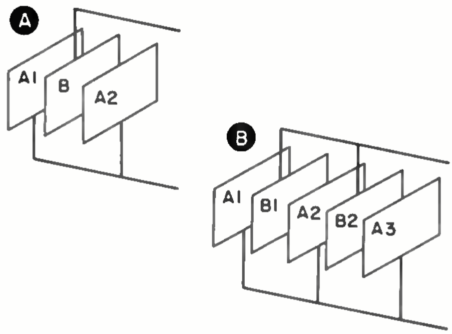


Fig. 3. Adding plates to a capacitor increases the capacity.

A1 opposite plate B is one square inch. When we add plate A2, we will have the areas of A1 and A2 exposed to both sides of B. Thus the effective area of the plates is doubled and therefore the capacity is doubled.

Additional plates can be added as shown in Fig. 3B. Adding additional plates to a capacitor actually increases the area of the plates, which in turn increases the capacity.

In considering the area of the plates, we must consider only the overlapping area. For example, if we have a capacitor made up of two plates each having an area of 1 square inch, and positioned as shown in Fig. 4A, we will have a certain capacity. However, if without changing the size of the plates we move them as shown in Fig. 4B, the capacity will be reduced, because the overlapping area of the plates is reduced. The part of plate A that is not directly opposite part of plate B will have little or no effect on the capacity. Similarly, the part of plate B that is not opposite part of plate A will have little or no effect.

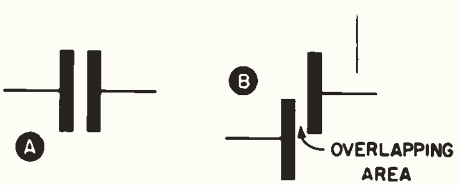


Fig. 4. Only the overlapping areas of the plates affect the capacity of a capacitor.

This is the principle that is used in variable capacitors. Here a number of plates are arranged so that one section of plates is movable and can be made to overlap more or less of the other section, thus exposing larger or smaller areas of the two sections to each other.

2. Spacing.

The distance between the plates of a capacitor will also affect the capacity. Capacity is inversely proportional to the spacing between plates. For example, if the spacing is reduced to one half, there will be twice the capacity. If it is reduced to one quarter, there will then be four times as much capacity. Also, if you double the space between the plates you will have one half of the capacity; if you triple the space you will have one third of the capacity. This is due to the fact that when the repelling effect of the electrons on one plate and the attracting force of the shortage of electrons on the other plate must act over a greater distance, they are not able to drive as many electrons out of the one plate of the capacitor and pull as many onto the other plate. As a result, the capacitor is not able to store as great a charge. You already know that the capacity of a capacitor is equal to the charge in coulombs divided by the voltage required to give that charge. Thus, if the charge a capacitor can hold is reduced, the capacity will go down. Moving the plates farther apart will reduce the charge that you can get on a capacitor for a given applied voltage. Conversely, bringing the plates closer together will increase the charge you can get on a capacitor for a given applied voltage.

3. Dielectric.

We have already mentioned that the type of dielectric used between the plates of the capacitor has an effect on the capacity. If instead of air between the plates we place a piece of mica, paper, or ceramic material, the capacity will be increased. As a matter of fact, if we slide a piece of mica between the plates of a capacitor so that the mica exactly fills up the space we will find that the capacity will increase somewhere between 6 and 8 times.

DIELECTRIC CONSTANT

In the preceding section we mentioned that inserting a piece of mica between the plates of a capacitor would increase the capacity somewhere between 6 and 8 times. The amount that a certain material will increase the capacity when used as a dielectric, compared to the capacity when air is used, is called the "dielectric constant" of the material. In other words, the dielectric constant tells us the number of times that the capacity will be increased by inserting a certain material between the plates of a capacitor. Actually, the dielectric constant is based on the number of times the capacity would be increased over the capacity we would have if the dielectric were a perfect vacuum. However, air between the plates affects the capacity very little. The capacity is practically the same with an air dielectric as it would be with a perfect vacuum between the plates so we say that the dielectric constant of air is 1; in other words it is the same as a vacuum between the two plates.

Different materials have different dielectric constants. Paper has a dielectric constant of somewhere between 1.5 and 3, depending upon the grade of paper. Mica has a dielectric constant between 6 and 8. Different types of oil have dielectric constants from about 2 up to about 10. The ceramic materials used in ceramic capacitors have a wide range of dielectric constants going up to as high as about 1500. You can see from this why it is possible to make ceramic capacitors with large capacities in small physical sizes.

We already mentioned that the effect of a dielectric is to increase the capacity of the capacitor. If there were a perfect vacuum between the plates of a capacitor, electrons flowing into one plate of the capacitor would place a negative charge on this plate and electrons flowing from the other plate would place a positive charge on it.

However when we place any dielectric material, which of course includes air, between the plates of the capacitor, the electrons on the negative plate of the capacitor distort the atoms of the material between the plates. Before the plates of the capacitor are charged, the atoms in the dielectric will be in their normal state with the electrons revolving around the nucleus as shown in Fig. 5A. However, when the electrons flow onto the negative plate of the capacitor, they will tend to repel the electrons in the dielectric. Since the dielectric is an insulator, the electrons are not free, but are bound to the atoms; however the repelling effect of the electrons on the negative plate will shift the path of the electrons in the dielectric so that the path around the nucleus will be like that shown in Fig. 5B. Notice that the electrons are pushed away from the negative plate and towards the positive plate. As these electrons move closer to the positive plate, they tend to repel the electrons

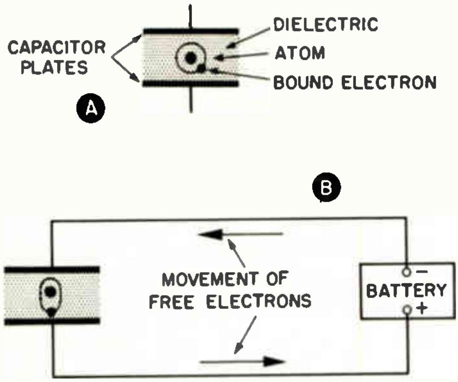


Fig. 5. Electrons in the dielectric are bound to their atoms, but their paths are shifted when a battery is connected across the capacitor, so that they come closer to the positive plate, thus transferring the effect of the extra electrons on the negative plate.

on the positive plate. Because the electrons from the atoms in the dielectric move over towards this plate of the capacitor, the dielectric has the effect of reducing the space between the plates. This places the negative charge very close to the positive plate and drives many more electrons off the positive plate of the capacitor than could be removed if the two plates were in a vacuum.

As we mentioned, air acts almost like a perfect vacuum. There is very little of this effect occurring when the dielectric is air. However, with materials of a higher dielectric constant this effect is more pronounced. This is particularly true in the case of ceramic materials where there is considerable distortion of the atoms in the dielectric so that the net effect is to get the same result as you would get by having the plates of the capacitor practically touching. Of course, in a practical case you can't have the plates practically touching, because they would probably short together and then the capacitor would be of no value. However, by using the right dielectric we get just as much capacity as we would with the plates practically touching and at the same time the dielectric between the plates holds the plates rigid and reduces the possibility of the plates accidentally shorting together.

VOLTAGE RATINGS

We mentioned previously that the charge that can be placed on a capacitor depends upon the electrical size or the capacity of the capacitor and also on the voltage used to place the charge on the capacitor. You might think from this that you could put more and more charge on a capacitor simply by increasing the voltage higher and higher. However, this is not the case because there is a limit to how much voltage can be applied to a capacitor.

Manufacturers design capacitors with a certain spacing between plates. If the plates are put very close together, you cannot put a very high voltage on the capacitor, because the electrons forced onto the one plate of the capacitor would jump right across the space between the plates to reach the other plate of the capacitor which has a shortage of electrons. Once this happens current will flow across the point where the capacitor breaks down, at least until you eliminate the short by shutting off the power. Sometimes when the electrons jump across the capacitor in this way the capacitor is permanently damaged.

Working Voltage.

When a manufacturer designs a capacitor he designs it for use in a circuit with a certain maximum operating voltage. The voltage is marked on the capacitor and is usually called the working voltage. The capacitors you will find in small pieces of electronic equipment such as radio and TV receivers will usually have a working voltage somewhere between 200 and 600 volts. Some of the capacitors in a TV receiver may have a working voltage as low as 200 volts. These capacitors are used in circuits where the operating voltage does not exceed 200 volts. Others used in higher voltage circuits may have a working voltage of 400 volts and still others used in circuits with an even higher voltage may have a working voltage of 600 volts or more.

One point to keep in mind when servicing electronic equipment is that the manufacturer of the equipment usually uses as low a working voltage as possible in order to keep the cost of the equipment low. However, there is no reason why a capacitor with a higher working voltage cannot be used providing there is room to do so.

Peak Voltage.

Sometimes you will find a capacitor with two voltage markings on it. It may be marked "working voltage 450 volts, peak voltage 525 volts". This type of marking is usually found on an electrolytic capacitor designed for use as a filter capacitor in a power supply. Here you know that the output of the rectifier in the power supply is pulsating dc. Thus the actual voltage at the output of the rectifier is not constant. If the dc output voltage from the power supply is 450 volts, during part of the time when the pulses from the rectifier tube reach their peak, the voltage will exceed this value. As long as this voltage peak does not exceed 525 volts, a capacitor marked with a 450-volt working voltage and a 525-volt peak voltage will work satisfactorily.

Some manufacturers do not mark electrolytic capacitors in this way. They simply mark them with the working voltage with the assumption that the peak voltage will not exceed a safe value.

SUMMARY

There are a number of important points that you should remember from this section of the lesson. First, remember that the basic action of a capacitor depends upon its ability to store an electric charge. Also remember that there is no complete circuit through a capacitor, but current will flow in a circuit in which a capacitor is connected while the capacitor is being charged and while it is being discharged. Remember that a capacitor is not charged instantly, but there is some time involved in charging a capacitor. The actual time it takes to charge a capacitor fully will depend upon the capacity and the resistance in the circuit.

Remember that the electrical size of a capacitor is measured in farads, but the farad is such a large unit that the practical values are the microfarad, which is a millionth of a farad, and the picofarad, which is a millionth of a microfarad.

The capacity of a capacitor depends upon the area of the plates, the spacing between the plates, and the dielectric between the plates.

The dielectric constant of the material is a number which tells you the number of times the capacity of a capacitor will be increased when this type of material is placed between the plates of the capacitor. Different materials have different dielectric constants; one of the highest is ceramic, which has a dielectric constant as high as 1500.

The voltage rating of a capacitor tells you the maximum safe voltage that you can apply to a capacitor. Capacitors having a higher voltage rating can always be used in replacing a defective capacitor in a piece of electronic equipment if there is room.

SELF-TEST QUESTIONS

(a) What do we mean when we say a capacitor is charged?

(b) What two factors affect the length of time it takes to charge a capacitor?

(c) Does a capacitor have any more electrons on its plates when it is charged than when it is discharged?

(d) What factor determines the amount of charge a capacitor can hold for a given applied voltage?

(e) What is the basic unit of capacity?

(f) What two practical units are used in electronics to indicate the capacity of a capacitor?

(g) Convert .0033 mfd to picofarads.

(h) Convert 680 pf to mfd.

(i) Name the three factors that affect the capacity of a capacitor.

(j) If you cut the spacing between the plates of a capacitor in half, will this double the capacity of the capacitor or will it cut the capacity of the capacitor in half?

(k) What is the dielectric constant of a material?

(l) What is the dielectric constant of air?

(m) What type of material has the highest dielectric constant?

(n) Why does a material with a dielectric constant greater than air have the effect of increasing the capacity of a capacitor?

(o) A capacitor is marked .02 microfarads, 400 volts; what does the marking 400 volts mean?

Typical Capacitors

Capacitors can be divided into two types, fixed capacitors and variable capacitors. A fixed capacitor is a capacitor which has a fixed capacity - in other words, the capacity of the capacitor cannot conveniently be varied. A variable capacitor, on the other hand, is a capacitor that is designed so that its capacity can be varied.

Capacitors can be further divided into types depending upon the dielectric material used. Most variable capacitors have an air dielectric and therefore are also referred to as air capacitors. Some variable capacitors have a mica dielectric and these are referred to as variable mica capacitors.

Most fixed capacitors use some dielectric other than air. Many capacitors use a paper dielectric and these are usually referred to as paper capacitors. Some capacitors of this type are dipped in a mylar coating to seal the capacitor and these are sometimes called paper mylar capacitors. Some fixed capacitors use a mica dielectric and these are called mica capacitors. Mica capacitors might also be sealed in a mylar case and these are sometimes called mylar-dipped mica capacitors.

Ceramic is widely used in capacitors as the dielectric and this type of capacitor is called a ceramic capacitor.

Large capacitors used in power supplies are called electrolytic capacitors. These capacitors use an electrolyte which forms a film that acts as the dielectric. You will study these capacitors in detail shortly and learn more about them at that time.

All capacitors, regardless of type, can store a charge. Of course, some capacitors can store more charge than others because they have a greater electrical size. In this section you will see that there is a reason for having these different types of capacitors, and also you will learn more about the common types of capacitors that you will encounter in electronics work.

VARIABLE CAPACITORS

A typical variable capacitor is shown in Fig. 6. This capacitor is actually two capacitors coupled by a common shaft, and as you know this type is called a two-gang capacitor. In some receivers you will find 3-gang capacitors, and in some old sets and in some communications

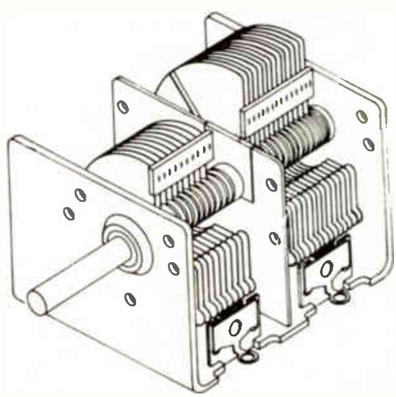


Fig. 6. A variable capacitor.

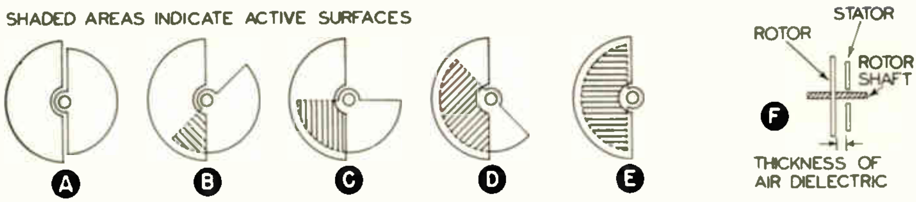


Fig. 7. How a variable capacitor with straight-line capacity plates works.

receivers you will find 4-gang capacitors. Sometimes all sections are identical; sometimes you will find that one section has smaller plates than the others.

Each section of a variable capacitor is made up of two sets of plates. The one set of plates that do not move are all connected together and insulated from the capacitor frame. These plates are called the stator plates, in other words, the stationary plates. The other set of plates are connected directly to the shaft and to the capacitor frame. These plates rotate and hence are called the rotor plates. Since these plates are connected directly to the capacitor shaft and to the frame, if you mount the capacitor on a metal chassis, the plates are automatically connected to the chassis. This does not present any problem, because in most circuits where a capacitor of this type is used, it is desirable to connect one set of plates to the chassis. The chassis acts as a ground or common connection for all rf (radio frequency) circuits.

There are two different types of plates found in variable capacitors. One type of plate is called the straight-line capacity type; the other is called the straight-line frequency type. An example of the straight-line capacity type of plate is shown in Fig. 7. Notice that in A of Fig. 7 the plates are completely separated so that the capacity is at a minimum. Actually there will be some capacity, because even though the plates are not meshed, the ends of the plates of the rotor have a certain capacity to the ends of the plates of the stator. In B the capacitor plates have been rotated through one eighth of a turn and the overlapping area of the two sets of plates is one quarter of the total area. We now have approximately one quarter of the total capacity. As the capacitor is turned another eighth of a turn, from position B to position C, so that it has completed a quarter turn, one-half of the area of the plates is overlapping and we have one half of the total capacity. Similarly, when the capacitor has been moved to three

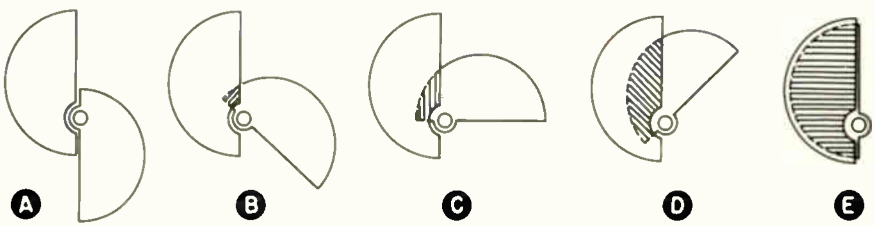


Fig. 8. How a variable capacitor with straight-line frequency plates works.

eighths of a turn, three-quarters of the area is overlapping and we have three quarters of the total capacity, and finally with a half turn the entire two areas are completely overlapping and we have maximum capacity.

There is another type of plate that is called the straight-line frequency type. Here the plates are shaped as shown in Fig. 8. Now when the plates are moved from position A to position B, less than a quarter of the area is overlapping even though the capacitor has gone through an eighth of a turn and we have less than a quarter of the total capacity.

Similarly when it is moved to position C, although the capacitor has been turned through a quarter of a turn, less than half of the two areas are overlapped and we will have less than half the total capacity. However, the increase in capacity obtained by rotating the capacitor from B to C is greater than was obtained in rotating it from A to B because the increase in overlapping area is greater. Similarly we get a still greater increase in capacity in going from C to D and an even greater increase in capacity in going from D to E.

The reason for this type of capacitor is that capacitors are used in conjunction with coils in the tuning circuits of radio and TV receivers, to select one station and reject others. For example, in a radio designed for the standard broadcast band you will find a coil and a capacitor used to select the desired station and reject the others. At the high end of the broadcast band it takes a smaller change in capacity to get a given frequency change than it does at the low end of the broadcast band. Therefore if you use the straight-line capacity type of capacitor, stations at the high-end frequency of the dial will be squeezed together and stations at the low-end frequency will be spread out. This makes it difficult to tune stations at the high frequency end of the dial. With the straight-line frequency type, however, the different frequencies are spread evenly

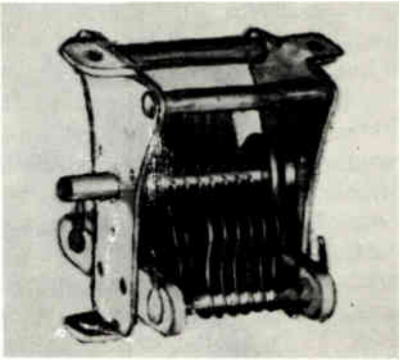


Fig. 9. A single-section variable capacitor of the transmitting type.

across the dial of the receiver so that it is just as easy to tune in a station at the high end of the dial as one at the low end of the dial. Most modern broadcast receivers use the straight-line frequency type of capacitor because it is somewhat easier to tune the receiver.

A single-section tuning capacitor such as might be found in a radio frequency transmitter or any other device where a variable capacitor is needed and high voltage is present is shown in Fig. 9. Notice that this capacitor is basically similar to one section of the capacitor shown in Fig. 6. The big difference is that the spacing between the plates is greater. The greater spacing is needed in high-voltage circuits to avoid arcing between the plates. Arcing is simply a flash-over that occurs where the voltage is so high that the electrons are able to jump from one plate of the capacitor to the other.

Trimmer Capacitors.

Trimmers are small variable capacitors, so called because they are used to trim or adjust resonant circuits whose main tuning capacitor is much larger.

Most ganged variable capacitors are equipped with trimmers so each ganged section can be individually adjusted. These trimmers use a mica dielectric, and the movable plate is of spring material which tends to stay away from the fixed plate, as shown in Fig. 10. A screw electrically insulated from the movable plate draws it closer to the fixed plate when tightened, thus increasing the capacity.

In some cases the trimmer is a miniature air dielectric capacitor, just like a single tuning capacitor.

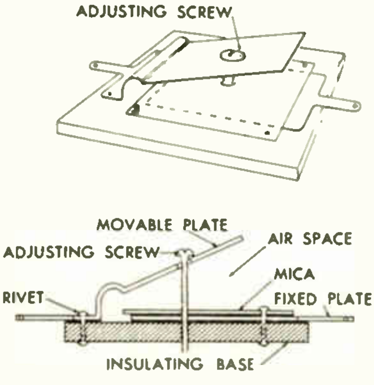


Fig. 10. A typical trimmer capacitor.

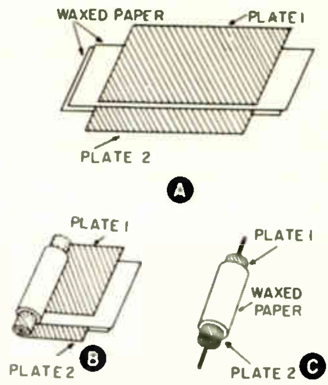


Fig. 11. How a paper capacitor is made.

PAPER CAPACITORS

A paper capacitor is made by taking two sheets of tinfoil and placing a sheet of paper between them as shown in Fig. 11A. The tinfoil and the paper are then rolled as shown in Fig. 11B until they are shaped like Fig. 11C. Wire leads are then attached to the foil sheets that protrude from each end of the capacitor.

After the leads have been attached to the capacitor it is then encased in a suitable container. In older radio and TV receivers you will find capacitors of this type that have been encased in a cardboard case with wax or some other sealing compound poured into both ends. However, this type of capacitor frequently caused trouble and is seldom used in modern electronic equipment. Modern paper capacitors are completely encased in a molded ceramic type of material or are encased in a mylar type of material. Either type completely seals the capacitor so that moisture cannot seep into it. Capacitors of this type are referred to as molded capacitors, molded paper capacitors, mylar capacitors or mylar paper capacitors.

A photo of two typical paper ca-

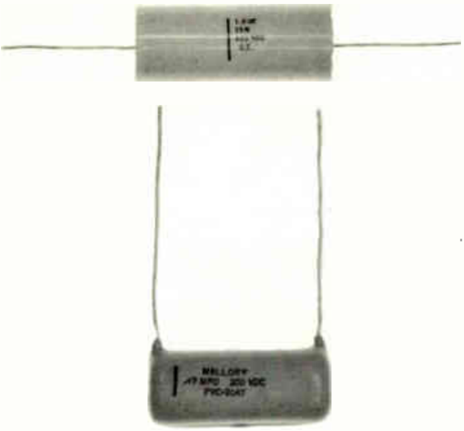


Fig. 12. Typical paper capacitors.

pacitors is shown in Fig. 12. Notice that in one capacitor the leads come out the ends. This type of lead arrangement is referred to as axial leads. In the other type of capacitor the leads come out the side and this is referred to as a radial-lead type. Capacitors with radial leads are frequently used in electronic equipment in which printed circuitry is used. This type of capacitor is very convenient for mounting in this type of construction. However, axial lead capacitors can also be used in the same application simply by bending the leads at right angles to the body of the capacitor.

Paper capacitors are made in a wide range of capacities. You will find paper capacitors as small as .0005 mfd and as large as 1 mfd or 2 mfd. Paper capacitors can be made in a larger size, but it is usually more economical to make other types when the capacity needed in a circuit exceeds about .5 mfd.

Defects.

Since you are training as an electronics technician, one of your chief concerns with capacitors will be the defects that occur in them. There are a number of different types of defects that can occur in paper capacitors. As we mentioned earlier, in older receivers you will find paper capacitors encased in a cardboard case. Moisture can seep into this type of capacitor, resulting in leakage from one plate to the other or an eventual breakdown in the paper insulation, so that for all practical purposes one plate is touching the other. Of course, when this happens the capacitor can no longer store a charge and acts as though there were a wire connected between the two leads of the capacitor.

Occasionally in a molded or a mylar type of capacitor the material encasing the capacitor will crack, particularly around the point where the leads are brought out. When this happens moisture can seep into this type of capacitor and you get exactly the same effect as moisture in the cardboard-encased capacitor.

It is usually not too difficult to identify a completely shorted capacitor, but one that has a high leakage may cause almost as much trouble as a shorted capacitor and it is much more difficult to find. There is no such thing as a perfect capacitor. There is some leakage between the plates of all capacitors. However, a good paper capacitor will usually have a leakage resistance of several thousand megohms. When the leakage resistance drops below this figure it is a sign that the capacitor is deteriorating. In some circuits a leakage resistance as low as 2 or 3 megohms can be tolerated and the circuit may work perfectly, but in other circuits a capacitor with a leakage resistance as high as 10 or 20 megohms may be totally unusable. It is too early for you to try to distinguish between these cases; the important thing for you to remember at this time is that leakage between the plates and direct shorts between the plates of capacitors are two defects that paper capacitors can develop.

Another defect found in paper capacitors is an open. The open is usually where one of the leads breaks loose from the tinfoil. Sometimes the lead will pull right out of the capacitor, and of course this is easy to spot because you can see it. But in most cases the break will be inside the capacitor so you cannot see it and you will have to rely on the way the circuit performs to give you a clue that this is a possible cause of trouble.

Another defect found in paper capacitors is an intermittent defect. Part of the time the capacitor will operate normally, but other times it may short or it may open. Again, we do not expect you to be able to find these defects at this time; the important thing is to remember that these types of defects can and do occur.

MICA CAPACITORS

Mica capacitors are somewhat larger physically than paper capacitors of equal capacity. However, mica is used in some capacitors that are to be used in high-frequency circuits. Mica capacitors are also used in the rf signal circuits of transmitters. Mica capacitors are made with capacities ranging from a few picofarads to approximately 10,000 pf. Some mica capacitors found in transmitters are designed for operation in circuits where the voltages are extremely high; capacitors with working voltages of 5000 volts are quite common.

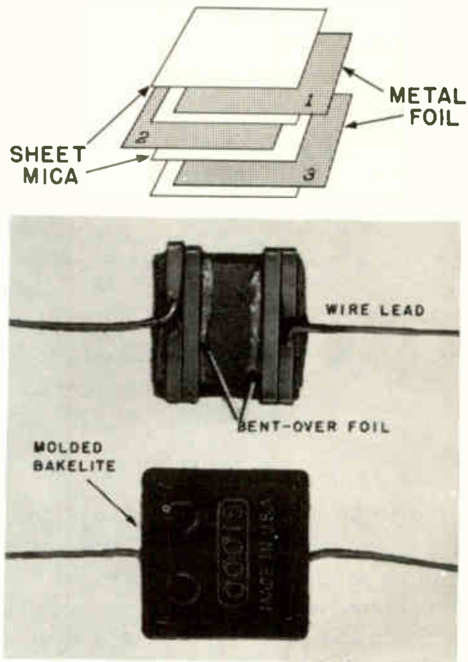


Fig. 13. How a mica capacitor is made.

The construction of a mica capacitor is shown in Fig. 13. Notice that there are simply two sets of plates, and that the plates are separated by thin sheets of mica placed between the metal plates. Because mica is brittle, this type of capacitor cannot be rolled into a tube like a paper capacitor and therefore most mica capacitors are rectangular in shape. The plates and the mica are enclosed in a ceramic or Bakelite case that is molded over the unit.

Mica capacitors seldom cause trouble. About the only defect that they ever develop is a short. Occasionally one of the leads will pull loose and the capacitor will open, but it is very rare to find any defect at all in a mica capacitor.

In spite of the fact that mica capacitors are almost trouble-free, they are not used too often in radio or TV receivers because they are more expensive than ceramic capacitors. However, in some critical circuits you will still find mica capacitors. You may find them used in the tuners of TV receivers or in critical sections of color TV receivers. They are still used almost exclusively in transmitting equipment.

As with paper capacitors, mica capacitors are available with both types of leads. Leads coming out of the end which are called axial leads are probably more common, but mica capacitors that have been dipped in mylar or some similar sealing compound are available with radial leads coming out of the side of the capacitor.

CERAMIC CAPACITORS

There are three types of ceramic capacitor found in electronic equipment. One type is the tubular, another is the disc. The third type is the feed-through. All three types are shown in Fig. 14.

Tubular ceramic capacitors are made in sizes ranging from less than 1 pf to about 1500 pf. They can be made to rather close tolerances and at one time were quite widely used in electronic equipment where small electrical size capacitors were needed. Most modern equipment, however, uses disc capacitors, because they are more economical than tubular capacitors and can also be made to quite close tolerances.

The feed-through capacitor is a special type of ceramic capacitor used exclusively as a bypass element, particularly at very high and ultra high frequencies. The unique construction of the feed-through capacitor, shown physically and schematically in Fig. 14, makes it a particularly effective bypass capacitor for filament, bias, agc and B+ leads in UHF and VHF TV tuners. One lead of the tubular feed-through capacitor connects directly to ground, whereas the fed through lead forms the other lead of the capacitor. The capacitor leads, therefore, are very short, making the capacitor extremely effective at UHF and VHF.

Disc capacitors are made with capacities from about 1 pf up to almost 1 mfd. These capacitors are made with different tolerances. Common tolerances are ±5%, or ±10% or ±20%.

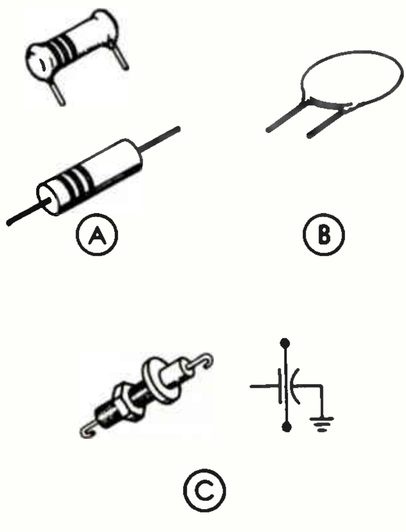


Fig. 14. Three types of ceramic capacitor. (A) Tubular; (B) Disc; (C) Feed-Through.

Sometimes the tolerance is indicated by a letter rather than by the actual tolerance figure. The letter J is used to represent a 5% tolerance, K a 10% tolerance and M a 20% tolerance. Ceramic capacitors are also made with a tolerance of +100% or -0. The letter P is used to represent this tolerance, and with a tolerance of +80%, -20% the letter Z is used.

Many types of ceramic capacitors change capacity with changes in temperature. For example, a capacitor that is labelled Z5U is a ceramic capacitor designed for operation between 10° and 85°C. The Z5 gives you this information. The letter U indicates that the capacity may change as much as +22% or -56% over that temperature range. In many circuits where ceramic capacitors are used a change in capacity is not particularly important and therefore you will find many capacitors used as bypass capacitors with the label Z5U.

Ceramic capacitors are made with lower temperature coefficients. For example, a capacitor labelled Z5F has a temperature coefficient of only ±7.5%. Z5P indicates a temperature change of 10% within the operating range.

When you have to replace a ceramic capacitor you can use one with the same temperature coefficient or one with a better temperature coefficient. The letter A indicates the smallest temperature coefficient, a change of ± 1%; the letter V is the poorest temperature coefficient. When you have to replace a ceramic capacitor, use one having the same temperature coefficient or a coefficient indicated by a letter closer to the letter A - this will be a better or closer tolerance temperature coefficient.

Ceramic capacitors normally have a voltage rating of at least 500 volts. Many have a voltage rating of 1000 volts. However, some will have a voltage rating as low as 100 volts where size is an important consideration. High-voltage ceramic capacitors having a voltage rating of 6000 volts or more are also available. Usually if the voltage rating is higher than 1000 volts or less than 500 volts, the voltage rating will be stamped on the capacitor.

Defects.

Occasionally a lead will break off a ceramic capacitor, but other than this they seldom open. Ceramic capacitors do short sometimes, but a shorted capacitor is usually not too difficult to locate. A low resistance reading across any capacitor indicates either that there is something directly across the capacitor causing the reading or else the capacitor itself is defective.

ELECTROLYTICS

There are two types of electrolytic capacitors, dry and wet. In the wet electrolytic capacitor a liquid called an electrolyte is used whereas in the dry electrolytic capacitor the electrolyte is in a paste form instead of a liquid. Wet electrolytic capacitors are no longer used. Modern electrolytic capacitors are all of the dry type.

A dry electrolytic capacitor is made of two plates with an electrolyte in paste form placed between the two plates as shown in Fig. 15. The anode plate is treated chemically before the capacitor is as-

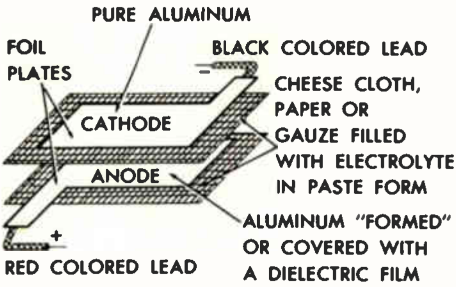


Fig. 15. How dry electrolytics are made.

sembled to produce a coating of oxide on the surface of the plate. The oxide then acts as the dielectric and the paste electrolyte acts as the other plate of the capacitor. The plate marked the cathode is the means of making contact to the paste.

Most electrolytics that you will encounter in commercial electronic equipment are made with aluminum plates. These capacitors are referred to as aluminum electrolytic capacitors. However, in some special applications, particularly where small size is important, the capacitor may be made with tantalum plates. This type of capacitor, however, is seldom found in radio or television receivers, because the tantalum plate electrolytic capacitor is much more expensive than an aluminum plate capacitor.

Polarity.

Electrolytic capacitors have polarity. This means that they can be used only in circuits having dc or pulsating dc. The plate called the anode must always be connected to the positive side of the voltage source and the plate called the cathode must always be connected to the negative side of the voltage source. If an electrolytic (technicians frequently shorten electrolytic capacitor to "electrolytic") is connected into the circuit backwards it will act as a low resistance, and a high current will flow through the capacitor, destroying it.

Dry electrolytic capacitors may be rolled into a tubular form and look very much like large paper capacitors. Flexible leads are brought out of the ends of the capacitor and the polarity of the leads is shown either by marking one end with a + sign and the other end with a - sign, or by using a wire of one color as one lead and a wire of another color as the other lead. The polarity is thus identified by the color of the lead. Usually a black wire is used to identify the negative lead and a red wire is used to identify the positive lead.

In some capacitors of this type there may actually be two electrolytic capacitors in the one container. The capacitor might be made with a common negative lead; in other words one negative lead for the two capacitors and separate positive leads so that there are only three leads coming from the capacitor or it can be made with separate positive and separate negative leads so that there are four leads brought out of the capacitor. The type with three leads, where a common negative lead is used, is found far more frequently than the type with the four separate leads.

Dry electrolytics are also sometimes placed inside a metal can. The can is the negative terminal of the capacitor and the positive terminal is brought up through an insulated wafer at the bottom of the capacitor. Sometimes there may be several separate capacitors in the one can. The can will be the common negative terminal and the separate positive leads will be brought out the bottom wafer. Symbols such as a small triangle, a half-moon or a square are cut into the wafer near the terminals to identify the various sections of the capacitor. An example of a can type electrolytic capacitor with four separate capacitors in the one container is shown in Fig. 16A. Fig. 16B shows a bottom view of the capacitor; you can see the symbols near the terminal lugs. The symbols identify the various terminals of the capacitor; the code used to identify the symbols is stamped into the metal capacitor can.

Defects.

As we mentioned previously, most electrolytic capacitors manufactured today are dry electrolytic capacitors. These capacitors deteriorate, particularly if they are not put into use. An electrolytic capacitor that has been unused for six months or more should be formed before the capacitor is put into service. An electrolytic can be formed by placing a low voltage on the capacitor and gradually increasing the voltage until it is equal to or slightly exceeds the rated working voltage of the capacitor. If a 450-volt capacitor that has been sitting around unused for six or seven months is simply installed in a circuit, and has a full 450 volts applied to it without first being formed, the chances are that it will short and the capacitor will be destroyed.

Electrolytic capacitors also deteriorate with use. The moisture in the electrolytic will slowly escape from

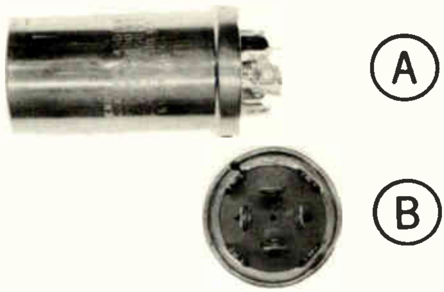


Fig. 16. A typical four-section electrolytic capacitor.

the capacitor and when all the moisture has escaped the capacitor will really be dry and it will no longer work. Electrolytic capacitors also develop leakage. Leakage in an electrolytic capacitor can be detected quite easily; you will notice that the capacitor starts to get hot. In normal operation there is some leakage through an electrolytic capacitor and this will cause the capacitor to get warm, but if you notice that an electrolytic is getting extremely hot, it is a sign that the leakage through the capacitor is too high, and the capacitor should be replaced.

SUMMARY

In this section of the lesson you learned that there are a number of different types of capacitors. There are two types of variable capacitors, those with an air dielectric and those with a mica dielectric such as compression-type trimmers. You also learned that there are paper, mica, ceramic, and electrolytic capacitors. You are likely to run into all types, but you will probably have more to do with electrolytics than the other types because electrolytic capacitors cause more trouble than

the others.

It is not important that you remember how the various types of capacitors are made; the important thing to remember is that capacitors can open, they can short, or they can develop intermittent defects. A low resistance reading across a capacitor indicates that the capacitor is shorted or has developed excessive leakage. Exactly how low a resistance can be tolerated through the capacitor depends upon the type of capacitor and the circuit in which it is used.

SELF-TEST QUESTIONS

(p) What two types of dielectrics are found in variable capacitors such as trimmer capacitors?

(q) What is the disadvantage of using a variable capacitor with straight-line capacity plates as a tuning capacitor in a radio receiver?

(r) Which type of capacitor plate is best suited for use in a tuning capacitor in a radio receiver?

(s) What is the main difference between a variable capacitor such as might be found in a radio receiver and a variable capacitor that might be found in a broadcast transmitter?

(t) What is the name given to the leads that come off the ends of a paper capacitor?

(u) Why are paper capacitors molded in a ceramic type of material or dipped in a mylar type of material?

(v) If the resistance of a paper capacitor is 100,000 ohms, would you replace the capacitor or is it satisfactory for continued use?

(w) Why are mica capacitors not more widely used in radio and television receiving equipment?

(x) In what type of circuits would you expect to find mica capacitors used?

(y) What two types of ceramic capacitors are used in electronic equipment?

(z) What does Z5F stamped on a ceramic disc capacitor indicate?

(aa) What are the two types of electrolytic capacitors, and which type is used in modern electronic equipment?

(ab) What do we mean when we say that an electrolytic capacitor has polarity?

(ac) If you notice that an electrolytic capacitor in a TV receiver is getting very hot, what should you do?

(ad) If you have a replacement electrolytic on hand that you have had for about a year, what should you do with it before installing it in a piece of equipment?

Capacitors in AC Circuits

Capacitors actually have very little use in circuits where there is nothing other than pure dc. Once a capacitor is placed in a dc circuit and charged, there will be no further current flow in the circuit. The chief importance of a capacitor comes from the way in which it works in ac circuits, and in circuits where there are ac and dc mixed together. Capacitors are used in all types of ac circuits found in electronic equipment ranging from the low frequencies found in power supplies and audio equipment up to the very high frequencies found in microwave equipment. Microwave equipment is that used at frequencies of 3000 mc (megacycles) and higher. The importance of the capacitor in ac circuits depends upon its ability to store an electrical charge. Because a capacitor can store a charge, it can be used in an ac circuit.

HOW AC FLOWS IN CIRCUITS USING CAPACITORS

In Fig. 17 a simple circuit is shown in which a capacitor is connected across an ac generator. In this type of circuit there will be a current flow. The exact amount of current flowing will depend upon the voltage of the generator, its frequency, and the capacity of the capacitor.

When the terminal of the generator marked 1 is negative and the terminal marked 2 is positive, electrons will flow from terminal 1 into the side of the capacitor marked A and force electrons out of the side marked B to terminal 2 of the generator which will attract these electrons because it is positive. During the next half cycle when the polarity of the generator reverses, electrons that have been piled up on the side of the capacitor marked A will be pulled out by terminal 1 which is positive, and electrons will be forced into side B of the capacitor by terminal 2 of the generator, which is negative. Extra electrons will be forced on this side of the capacitor so the side marked B will become negative and side marked A will become positive.

This action continues as the generator goes through first one half cycle and then the other. Electrons will flow back and forth in the circuit. They will flow first into one side of the capacitor and force electrons out of the other side and then electrons will flow out of the side on which they built up a surplus and into the side on which there was a shortage. It is important that you notice that electrons do not flow through the capacitor. You will remember that the plates of the capacitor are separated by a dielectric and the dielectric is a non-conducting material. However, because the capacitor can store a charge, we have the effect of a current flowing in the circuit.

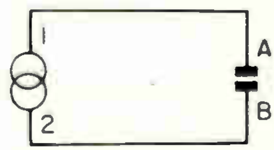


Fig. 17. A capacitor connected across a generator.

The action inside the capacitor can be seen in more detail in Fig. 18. At the start of the ac cycle when the voltage of the generator is zero, there will be a certain number of electrons on both plates of the capacitor. The electrons in each atom of the dielectric will be revolving around the nucleus as shown in Fig. 18A. However, when electrons begin to move into one plate, as shown in Fig. 18B, and out of the other plate, the electrons in the dielectric will be forced out of their normal path as shown. Thus, although the electron flowing into plate 1 does not reach plate 2, it does force another electron in the dielectric over near plate 2, and this in turn forces an electron out of plate 2.

As the ac voltage decreases and finally drops to zero, the electrons in the dielectric will return to the normal position as shown in Fig. 18C. During the next half cycle, when the polarity of the generator reverses, electrons will be forced into plate 2 and they in turn will force the electrons in the dielectric out of their normal positions and they will push electrons out of plate 1. Thus we have the effect of current flowing through the capacitor in the opposite direction, although the electrons flowing into the one plate never do get through the dielectric into the other plate of the capacitor.

You can see that there is a back and forth motion of the electrons in the conductors connected to the capacitor. The electrons in the dielectric will move back and forth and therefore we are justified in saying that ac current flows "through" a capacitor, even though the electrons never get through the dielectric into the other plate. Because of this effect, capacitors can be used in ac circuits. They are very useful in circuits where we have both ac and dc. The capacitor can be used to block dc, while at the same time allowing ac to flow through the capacitor.

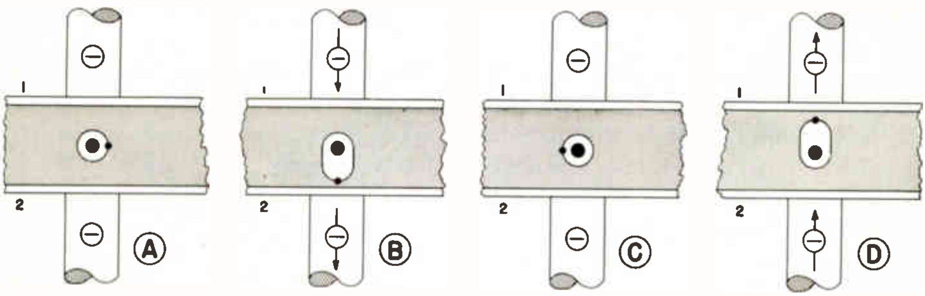


Fig. 18. When ac is applied to a capacitor, the bound electrons in the dielectric move first one way, then the other, so in effect, alternating current flows through the capacitor.

The action of a capacitor in allowing electrons to flow back and forth is a good demonstration of what ac is. The actual movement of each electron is very small; however, there may be a large number of electrons moving back and forth over a very short distance. The distance of the electron's travel is unimportant; the important thing is the number of electrons in motion. If we have a large number of electrons in motion, we have a large current.

The capacitor does not allow electrons to move back and forth without offering opposition. Capacitors do offer opposition to the flow of ac current through them and this opposition is called capacitive reactance.

CAPACITIVE REACTANCE

Since work must be done to move the electrons in the dielectric back and forth to permit ac current to flow, in a capacitive circuit there is opposition to the flow of current. This opposition is called capacitive reactance. This opposition is measured in ohms, just as the inductive reactance of a coil is measured in ohms. However, there is a great deal of difference between inductive reactance and capacitive reactance.

Capacitive reactance is represented by the symbol Xc. It can be expressed by the formula:

$$Xc = \frac{1}{6.28 \times f \times C}$$

In this formula the frequency f is the frequency expressed in cycles and C is the capacity in farads. We can write this formula in another way by expressing C in microfarads. To do this, we divide 6.28 into 1 and multiply this result by 1,000,000. We get:

$$Xc = \frac{159{,}000}{f \times C}$$

In this expression f is the frequency in cycles per second, and C is the capacity in microfarads.

From this formula there are several important things that you can see. First of all, let us consider the effect of a change in frequency on the reactance of a capacitor. Let us find the reactance of a 1-mfd capacitor at a frequency of 10 cycles per second. Using the formula:

$$Xc = \frac{159{,}000}{f \times C}$$

and substituting 10 for f and 1 for C we get:

$$Xc = \frac{159{,}000}{10 \times 1} = 15{,}900 \text{ ohms}$$

When the frequency is 100 cycles, we get:

$$Xc = \frac{159{,}000}{100 \times 1} = 1{,}590 \text{ ohms}$$

Notice that at the higher frequency, the capacitive reactance is lower.

As the frequency increases, the capacitive reactance decreases. In an inductive circuit we had just the opposite effect; if the frequency increased, the inductive reactance increased.

We have the same situation when the capacity is increased. If the capacity is made larger, the capacitive reactance decreases. We can see this if we find the reactance of a 1-mfd capacitor at a frequency of 100 cycles per second and then find the reactance of a 10-mfd capacitor at 100 cycles per second. We already know that the 1-mfd capacitor has a reactance of 1590 ohms. To find the reactance of the 10-mfd capacitor we use:

$$Xc = \frac{159{,}000}{100 \times 10} = 159 \text{ ohms}$$

Notice that this is one tenth the reactance of the 1-mfd capacitor at 100 cycles per second. Therefore in a capacitive circuit, we have exactly the opposite effect to what we had in an inductive circuit. We can say that the capacitive reactance varies inversely with the frequency and the capacity. This simply means that if the frequency or capacity increases, the reactance decreases, and if the frequency or capacity decreases, the capacitive reactance increases.

While we are discussing capacitive reactance it might be well to point out that the reactance of even a small capacitor becomes quite small if the frequency is made high enough. For example, a 100-pf capacitor has a reactance of 1590 ohms at a frequency of 1 megacycle. One megacycle is not a high radio frequency; as a matter of fact this frequency falls in about the middle of the standard radio broadcast band. At a frequency of 10 megacycles, which is in the short-wave bands, the reactance is only 159 ohms, and at a frequency of 100 megacycles, which is in the FM broadcast band, the reactance is only 15.9 ohms. Even a 1-mmf capacitor has a reactance of only 1590 ohms at a frequency of 100 megacycles. Thus, if the frequency is made high enough even small capacitors have a comparatively low reactance.

Capacitors in Parallel.

When two capacitors are connected in parallel we have two capacitive reactances in parallel. If the capacitive reactance of each capacitor is 100 ohms, we have the same effect as we would have with two 100-ohm resistors in parallel. The reactance would be only 50 ohms. This effect is the same as connecting a capacitor twice as large as either capacitor into the circuit.

To find the total capacity of capacitors connected in parallel you simply add the capacities. In other words, if a 4-mfd capacitor is connected in parallel with a 6-mfd capacitor, the total capacity in the circuit is 10 mfd. The reactance in the circuit would be exactly the same as you would obtain by connecting a 10-mfd capacitor in the circuit.

This is an important rule to remember--to find the total capacity of parallel-connected capacitors you simply add the capacitors together. The working voltage that can be applied to the parallel combination is the lowest working voltage of the capacitors connected in parallel.

Capacitors in Series.

When capacitors are connected in series, we have two reactances in series. The total capacitive reactance in the circuit is equal to the sum of the two reactances, just as the total resistance in a circuit made up of resistors connected in series is equal to the total resistance in the circuit. Thus connecting capacitors in series increases the total reactance in the circuit. If the reactance in the circuit increases, then the capacity must decrease.

When two capacitors are connected in series, you can find the total capacity by using the formula:

$$C_T = \frac{C1 \times C2}{C1 + C2}$$

When three or more are in series, use the formula:

$$C_T = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots}$$

When capacitors are connected in series, the total capacity is always less than the capacity of the smallest capacitor.

You will seldom actually have to calculate the value of series and parallel-connected capacitors. However, it is important for you to realize that connecting capacitors in series results in a lower total capacity, while connecting them in parallel results in a higher total capacity in the circuit.

SUMMARY

There are several important things that you should remember from this section of the lesson. First, remember that although current does not actually flow through a capacitor, the effect produced in the circuit is the same as though current does flow through the capacitor. Thus we say that an ac current flows through a capacitor.

Remember that we call the opposition offered to current flow in an ac circuit by a capacitor the capacitive reactance, and the capacitive reactance is measured in ohms. The capacitive reactance is equal to:

$$X_C = \frac{159{,}000}{f \times C}$$

where f is in cycles per second and C is in microfarads.

Increasing the capacity or the frequency in the circuit will result in a lower capacitive reactance and decreasing the frequency or capacity will result in a higher capacitive reactance.

When capacitors are connected in parallel, the total capacity is equal to the sum of the capacities. When capacitors are connected in series, the total capacity is always less than the capacity of the smallest capacitor.

SELF-TEST QUESTIONS

(ae) When a capacitor is connected across an ac generator, does current flow from the generator?

(af) When a capacitor is connected across a generator, does current flow through the capacitor?

(ag) What is the name given to the opposition that a capacitor offers to the flow of ac?

(ah) What is the formula used to determine the capacitive reactance of a capacitor?

(ai) What is the capacitive reactance of a .02 mfd capacitor at a frequency of 150 cycles per second?

(aj) What is the total capacitance of a .1 mfd and a .22 mfd capacitor connected in parallel?

(ak) A 27 pf capacitor and a 56 pf capacitor are connected in series. What is the total capacity of the series-connected capacitors?

(al) In what unit is capacitive reactance measured?

(am) Complete the following statement: Increasing the capacity or the frequency in a circuit will result in a ________ capacitive reactance and a decrease in frequency or capacity will result in a ________ capacitive reactance.

Simple RC Circuits

An RC circuit is a circuit containing resistance and capacity. These circuits are found in all types of electronic equipment. RC circuits are used to shape signals. We can apply a signal with one type of wave shape to an RC circuit and get a signal having a different wave shape at the output. RC circuits are used to feed signals from one stage to another in electronic equipment. There are many applications of RC circuits, but before you can understand how these circuits are used, you must learn something about the fundamentals of the circuit.

You already know that when a capacitor is connected across a voltage source it does not charge instantly, but takes a certain length of time to charge. The length of time depends upon the size of the capacitor and the resistance in the circuit. The length of time it takes to charge up to a certain value is called the time constant. Time constant is an important consideration in circuits using resistance and capacitance. Let's learn a little more about it.

TIME-CONSTANT

When a resistor is connected in series with a capacitor and the two are connected across a battery as shown in Fig. 19, current begins to flow in the circuit to charge the capacitor. At the first instant that the resistor and capacitor are connected across the battery, a rather large current flows because there is no charge on the capacitor. The size of the resistor will limit the amount of current that can flow for any given voltage source. At this instant, although the current flowing in the circuit is high, the voltage across the capacitor is zero. In other words, when the current flowing into the capacitor is at a maximum, there is no voltage across it.

Gradually the capacitor is charged and as the capacitor charges, the voltage across the capacitor builds up and the current flowing in the circuit decreases. This is due to the fact that the actual voltage driving electrons in the circuit is equal to the source voltage minus the voltage across the capacitor. As the voltage across the capacitor increases, the voltage forcing electrons through the circuit goes down. In other words the current flowing into the capacitor decreases as its voltage increases.

If we draw a graph showing the way in which a capacitor charges, it would look like Fig. 20. In this graph, time is measured along the horizontal axis. The extreme left of this axis represents the instant we connect the capacitor and resistor across the source. Notice that at this instant there is no voltage across the capacitor, but the voltage starts

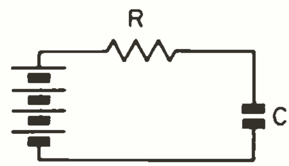


Fig. 19. A resistor and a capacitor connected in series across a battery.

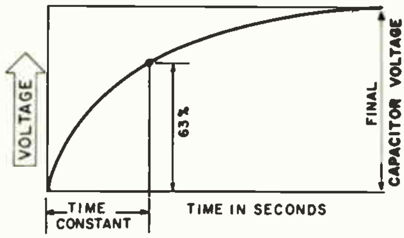


Fig. 20. How a capacitor charges. The time-constant is the time it takes for the capacitor to charge to 63% of the total voltage.

to build up rapidly. Then, as the charge on the capacitor increases, the rate of charge decreases so that when the capacitor is almost charged to a value equal to the source voltage, the rate of charge becomes very slow.

When we speak of the time-constant of an RC circuit, we mean the time it takes the capacitor to charge up to about 63% of the total source voltage. This value is shown on Fig. 20. The time-constant of any RC circuit in seconds can be found by multiplying the resistance of the resistor in megohms times the capacity of the capacitor in mfd. Thus, if a 2-mfd capacitor is connected in series with a 1-megohm resistor, the time-constant will be 2 × 1 = 2 seconds. This means that it will take two seconds for the capacitor to charge up to 63% of the source voltage.

One important thing to note is that the source voltage has nothing to do with the time-constant of the circuit. In other words, whether a resistor-capacitor combination is connected across a 10-volt battery or a 100-volt battery, the time-constant is the same. If a certain resistor and capacitor are connected across a 10-volt battery and the time-constant is one second, the capacitor will charge up to 63% of 10 volts, or 6.3 volts in one second. On the other hand, if the same resistor-capacitor combination is connected across a 100-volt battery, the capacitor will charge up to 63% of 100 volts, or 63 volts, in one second. The voltage across which the resistor-capacitor combination is connected does not determine the time-constant; the only factors that affect the time-constant are the resistor and capacitor values.

If a .05-mfd capacitor is connected in series with a 100,000-ohm resistor, the time-constant will be .05 × .1 (100,000 ohms = .1 meg) = .005 second. Thus, you can see that the time-constant of the combination using a smaller capacitor and resistor is much shorter than the time-constant of the combination of the two-mfd capacitor and the one-megohm resistor. Decreasing the size of either the resistor or the capacitor decreases the time-constant of the circuit, and increasing the size of either the resistor or the capacitor increases the time-constant of the circuit.

VOLTAGE-CURRENT PHASE

In the preceding example, you will notice that when the capacitor and resistor combination is first connected across the battery, a very high current flows in the circuit. However, at this first instant when they are connected across the bat-

 $T_C = RC$

R = MEG C = MFD

tery, there's no voltage across the capacitor. Here we have a situation where the current is at a maximum, and the voltage across the capacitor is at a minimum or zero voltage.

As the voltage across the capacitor builds up, the current flowing in the circuit decreases until finally, when the capacitor is fully charged to a value equal to the battery voltage, the current drops to zero, because the source voltage is unable to force any additional electrons onto the one plate of the capacitor or pull electrons off the other plate.

Essentially the same situation exists when a capacitor is used in an ac circuit. When the ac voltage across the capacitor builds up, the current decreases until, by the time the capacitor is fully charged, the current has dropped to zero. When the voltage across the capacitor begins to decrease, then current must flow in the opposite direction in order for the electrons to remove the excess electrons from the plate of the capacitor having a surplus of electrons and to replace the missing electrons on the plate of the capacitor having a shortage of electrons. During the second half of an ac cycle, as the capacitor voltage drops from a maximum value to zero, the discharge current flowing in the circuit increases until, at the time the voltage reaches zero and begins to change polarity, the current flowing in the circuit is at maximum. As the voltage across the capacitor builds up in the opposite direction, the current begins to decrease until, at the instant when the capacitor is fully charged with the opposite polarity, the current flowing in the circuit has dropped to zero again.

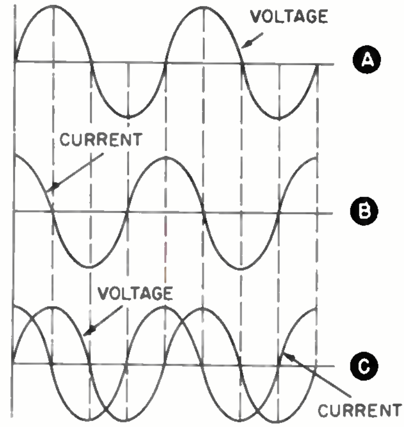


Fig. 21. The phase relationship between current and voltage in a capacitive circuit.

In Fig. 21 the relationship between the current and voltage in a capacitive ac circuit is shown. Notice that in Fig. 21A two cycles of an ac sine wave are drawn. Immediately beneath the voltage sine wave, in Fig. 21B, the current that will flow at the corresponding instant is shown. The voltage and current waves are superimposed in Fig. 21C.

In this circuit the effect is exactly opposite to that obtained when a coil is used. You will remember that in inductive circuits the current lags behind the voltage. Here, in a capacitive circuit, the current is leading the voltage. In a circuit containing only capacity, the current will lead the voltage by 90 degrees. In a circuit containing both resistance and capacity, the current will lead the voltage by a value somewhat less than 90 degrees; the actual phase difference between the voltage and current will depend upon the size of the capacitor and resistor in the circuit.

Thus, you have now seen two important examples of phase. In an inductive circuit, the current lags the voltage by 90 degrees. In a capacitive circuit, the current leads the voltage by 90 degrees.

VOLTAGE DISTRIBUTION

If we connect a capacitor in series with a resistor and connect the two across a dc voltage source, we will eventually have the situation shown in Fig. 22A. Here, all the voltage appears across the capacitor and there is no voltage across the resistor. Since there is no voltage across the resistor, we know immediately that there is no current flowing in the circuit. Of course, you know that at the instant the resistor and capacitor in series are connected across the voltage source, there will be a current flow while the capacitor is charging. However, once the capacitor is charged, there is no further current flow in the circuit and the meters would read like those shown in Fig. 22A.

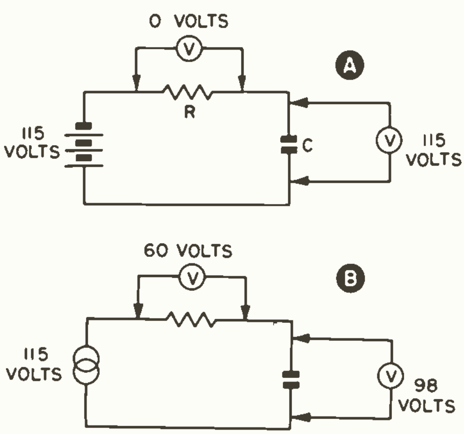


Fig. 22. When a resistor and a capacitor are connected in series across a dc source as at A, there is no current flow in the circuit. When the resistor and capacitor are connected in series across an ac generator as at B, there is current flowing, and hence there are voltage drops across R and C.

If the same resistor and capacitor are connected across a 115-volt ac source and the meters are replaced by ac voltmeters, we might encounter the situation shown in Fig. 22B. Here we have 98 volts across the capacitor and 60 volts across the resistor. 98 plus 60 adds up to 158 volts, which is more than the source voltage. It is obvious that we cannot add these two voltages in order to get the source voltage, because we know that the sum of the voltage drops in a circuit must be equal to the source voltage. The reason for the apparent contradiction is that the voltmeters indicate the RMS or effective voltage appearing across the capacitor and across the resistor. These voltages are not in phase and hence cannot be added by simple addition. You will remember that we encountered exactly the same thing when a resistor was connected in series with a coil. We found that we could not simply add the voltages appearing across the two and get the source voltage, but instead had to add the two by means of vectors. Let us see how we can do the same thing with these two voltages.

First, we start by drawing the current vector as shown in Fig. 23A. This vector is always drawn in this position since we use this as a starting point. We consider this vector as rotating in a counterclockwise direction around its starting point as the current goes through its cycle. If we use a scale of 50 volts to an

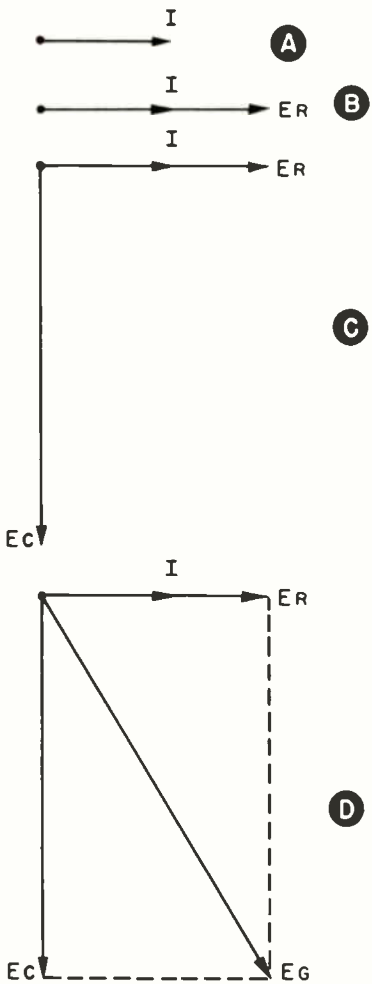


Fig. 23. Vector addition of two voltages.

inch, we can draw the vector E_R, representing the voltage across the resistor, 1-1/5 inches long as shown in Fig. 23B. We know that this vector must be drawn right on top of the current vector because the voltage across the resistor is in phase with the current flowing through it.

The next vector to draw is the vector representing the voltage appearing across the capacitor. We know that in a capacitor the current leads the voltage by 90 degrees. Another way of saying this is that the voltage lags the current by 90 degrees. Since the current vector is rotating in a counterclockwise direction, to show the voltage vector 90 degrees behind the current vector, we draw it as shown in Fig. 23C. Since the voltage across the capacitor is almost 100 volts, we can draw this vector 2 inches long as shown. Now we have only to draw in the dotted lines as shown in Fig. 23D and complete the vector diagram by drawing in the vector E_G which is the vector sum of the two voltages. If you measure this vector, you will find that it is a little more than 2-1/5 inches long, indicating a voltage of about 115 volts. In other words, when we used this scheme of adding the two voltages, we found that the voltage across the capacitor plus the voltage across the resistor, when added by means of vectors, are equal to the source voltage.

Another method of arriving at the same result is by means of the formula:

$$E_T = \sqrt{E_R^2 + E_C^2}$$

To solve this we take E_R^2, which is $60 \times 60 = 3600$, and add this to E_C^2 which is $98 \times 98 = 9604$. The sum is 13,204. $E_T = \sqrt{13{,}204}$, and the square root of 13,204 is approximately 115. Therefore the sum of the two voltages is 115 volts. As we pointed out before, if you know how to handle squares and square roots, the mathematical solution is somewhat quicker but if you do not know how to do square root, then the vector solution is entirely satisfactory.

IMPEDANCE

Because the voltage that will appear across each component in an RC circuit will depend upon the resistance or reactance of the particular part, we can draw impedance diagrams to obtain the total impedance in the circuit using the same procedure we used to add the voltage in a circuit. As an example, suppose we have a 1000-ohm resistor connected in series with a capacitor having a reactance of 1000 ohms. If we want to find the total impedance in the circuit, we proceed as follows:

First, draw the current vector as shown in Fig. 24A. Since the voltage appearing across the resistor will be in phase with the current flowing, we draw a resistance voltage vector immediately on top of the current vector. The voltage across the resistor will depend on its resistance. If we use a scale of 1000 ohms equals 1 inch, the resistance vector R is drawn, as shown in Fig. 24B, 1 inch long.

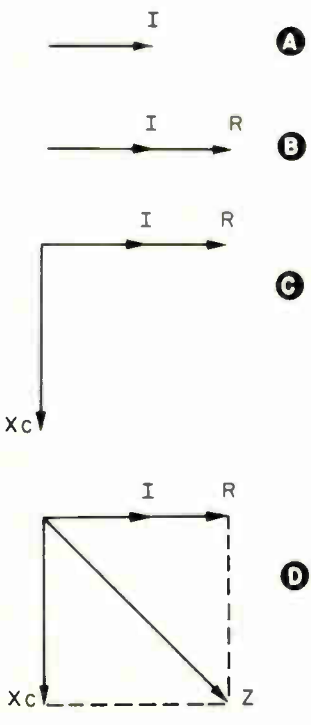


Fig. 24. Vector addition of resistance and capacitive reactance to find impedance.

Next, draw the vector shown in Fig. 24C to represent X_C. This vector also should be 1 inch long. It is drawn as shown, because the voltage appearing across the capacitive reactance will lag the current flowing in the circuit by 90 degrees.

We now complete the impedance diagram as shown in Fig. 24D, and you will find that the impedance vector Z will be 1.41 inches long. Since the scale we used was 1000 ohms per inch, the impedance of the circuit is 1410 ohms.

The same result could have been obtained mathematically by squaring the resistance and the capacitive reactance and adding the two together and taking the square root of the sum as indicated in the formula:

$$Z = \sqrt{R^2 + X_C^{\ 2}}$$

It is important to notice the difference between the capacitive reactance and the inductive reactance. Notice that one is simply the opposite of the other. Notice that the impedance diagrams are drawn differently. In Fig. 25A, we see the solution of a circuit using a scale of 1000 ohms per inch where a 1000-ohm resistor is connected in series with the capacitor having a reactance of 1000 ohms. In Fig. 25B we

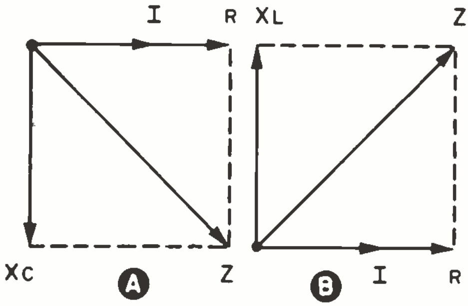


Fig. 25. When resistance and capacitive reactance are added as at A, the impedance vector lags the current; but when resistance and inductive reactance are added as at B, the impedance vector leads the current.

see an impedance diagram using the same scale, where a 1000-ohm resistor is connected in series with a coil having a reactance of 1000 ohms. Notice that we end up with the same impedance, 1410 ohms, in each case, but also notice the fact that in one case, the impedance vector leads the current vector, whereas in the other case, the impedance vector lags the current vector. In a later lesson, you will see more about these circuits and will also see the effect of having capacitive reactance and inductive reactance in the same circuit.

SUMMARY

In this section of the lesson you began the study of a very important phase of your course; you began to see what happens when resistance and capacitance are both used in the same circuit. You found that it takes a certain length of time for a capacitor to charge when the two are used in a dc circuit, and the length of time depends upon the size of the capacitor and the resistance in the circuit. The time-constant of a circuit is equal to the product of the resistance in megohms times the capacity of the capacitor in microfarads. The time-constant of the circuit is the length of time it would take to charge the capacitor to a value of 63% of the total voltage applied to the circuit.

You learned that in an ac circuit the current flowing in a capacitive circuit will lead the voltage by 90 degrees. In other words, a capacitive circuit acts exactly the opposite to an inductive circuit.

In an ac series circuit with a resistor and capacitor in series the voltage appearing across the resistance is not in phase with the voltage across the capacitor. If we have to add these voltages together, we must add them by means of vectors.

You also have learned that the impedance in an ac circuit using a resistor and a capacitor is the total opposition to current flow. The impedance can be obtained by means of the vector addition of the resistance plus the capacitive reactance in the circuit.

SELF-TEST QUESTIONS

(an) What do we mean by the time-constant of an rc circuit?

(ao) What is the time-constant of a .02 mfd capacitor charging through a 2.2 meg resistor?

(ap) If the time-constant of an rc circuit is 2 seconds when it is charged across a 10-volt battery, what will the time-constant be when it is connected across a 100-volt battery?

(aq) When a capacitor is connected across an ac generator, what is the phase relationship between the voltage and current in the circuit?

(ar) If a resistor and a capacitor are connected in series across an ac generator, what will the phase relationship between the voltage and current be?

(as) If a generator is connected across a resistor and a capacitor in series, what is the generator voltage if the voltage across the resistor is 9 volts and the voltage across the capacitor is 12 volts?

(at) What do we mean by the impedance in a series rc circuit in which a resistor is connected in series with a capacitor?

(au) If a 6-ohm resistance is connected in series with a capacitor having a capacitive reactance of 8 ohms, what will the impedance of the combination be?

LOOKING AHEAD

Up to this point in your course, you have been studying the basic action of a few important components found in electronic circuits. Other than tubes and transistors, the parts you will run into most frequently are resistors, coils, and capacitors or parts made of these three basic components. Now that you have studied each of these three parts separately, you are in a position to go ahead to see how they work together. A later lesson will discuss resonance. Resonance is perhaps one of the most important things you will study, because if it were not for resonant circuits many of the electronic miracles that we have today would not be possible.

ANSWERS TO SELF-TEST QUESTIONS

(a) When we say a capacitor is charged, we mean there is a surplus of electrons on one plate and a shortage of electrons on the other. When a capacitor is completely charged by connecting it across a voltage source, the voltage existing across the plates of the capacitor will be equal to the source voltage used to charge it.

(b) The length of time it takes to charge a capacitor is affected by the resistance in the circuit and the capacity of the capacitor.

(c) A charged capacitor does not have any more electrons on its plates than when it is discharged. Although the number of electrons on the two sets of plates is the same, some electrons have been removed from one plate and added to the other when the capacitor is charged.

(d) The capacity of the capacitor determines the amount of charge a capacitor can hold for a given applied voltage.

(e) The farad is the basic unit of capacity.

(f) The microfarad and the picofarad are the two practical units used in electronics. A microfarad is equal to one millionth of a farad, and the picofarad is equal to one millionth of a microfarad.

(g) 3300 pf. To convert microfarads to picofarads, you multiply by 1,000,000 or move the decimal point six places to the right.

(h) .00068 mfd. To convert picofarads to microfarads you divide by 1,000,000 or move the decimal point six places to the left.

(i) (1) the area of the plates (2) the spacing between the plates (3) the dielectric of the medium between the plates.

(j) The capacity will be doubled. The capacity of a capacitor is inversely proportional to the spacing between the plates. If you cut the spacing in half, the capacity will be doubled, but on the other hand if you double the spacing between the plates, the capacity will be cut in half.

(k) The dielectric constant of a material tells you how many times the capacity of the capacitor will be increased by substituting the material between the plates of the capacitor in place of air. In other words, if a material has a dielectric constant of 5, if you fill the spacing between the plates completely with that material, then the capacity will be increased five times from what it would be if the dielectric was air.

(l) The dielectric constant of air is 1.

(m) Ceramic-type materials have the highest dielectric. This is why ceramic capacitors of a given capacity will be smaller than any other type of capacitor.

(n) A material with a dielectric constant greater than air increases the capacity because it has the effect of reducing the spacing between the plates of the capacitor.

(o) The marking, 400 volts, on the capacitor means that the capacitor can be used in circuits having a dc voltage up to 400 volts. In other words, the rated maximum voltage at which the capacitor can be used is 400 volts. If you use a capacitor in a circuit where the dc voltage is 500 or 600 volts, the chances are it will break down in a short time. On the other hand a capacitor rated at 400 volts can be used in a circuit where the voltage is 300 volts or any value less than 400 volts.

(p) Air and mica.

(q) The stations on a high end of the band will be very close together and difficult to separate. On the other hand, the stations on the low end of the band will be spread out much more than they need be.

(r) A capacitor with straight-line frequency plates.

(s) The transmitter capacitor will have a much greater spacing between the plates than the receiver capacitor. The greater spacing is required to prevent arc over due to the high voltages used in the transmitter.

(t) Axial leads.

(u) To seal the capacitor so moisture can't seep into the capacitor and cause it to break.

(v) The capacitor should be replaced; a leakage resistance of 100,000 ohms is much too low for a paper capacitor.

(w) There are more economical types available which are almost as good as mica capacitors.

(x) In critical circuits where extremely stable capacitors are required.

(y) Tubular ceramic capacitors and disc type ceramic capacitors.

(z) Z5 indicates that the capacitor is designed for operation between 10° C and 85° C. The F indicates that the capacity will not change more than ±7.5% within the capacitor's normal operating range.

(aa) The two types of electrolytic capacitors are the wet and dry type. The dry type is used in modern electronic equipment; the wet type of capacitor is no longer made.

(ab) When we say a capacitor has polarity it means that it is made so that one plate must always be connected to a positive voltage and the other always to the negative voltage.

(ac) You should replace the electrolytic capacitor. The fact that it is getting very hot indicates that there is excessive current flow through it due to excessive leakage.

(ad) The capacitor should be formed. You can form it by placing a low voltage on it and gradually increasing the voltage until it is equal to or slightly exceeds the rated voltage of the capacitor. If you do not reform the capacitor and simply install it in the receiver, the chances are that it will break down.

(ae) Yes. Current flows from the generator to charge the capacitor with one polarity during one-half cycle, and then flows in the opposite direction to charge the capacitor with the opposite polarity during the next half cycle.

(af) No. Electrons flow into one plate and out of the other to charge the capacitor with one polarity during one half cycle, and then electrons flow out of the first plate and into the second to charge the capacitor with the opposite polarity during the second half cycle. Current flows back and forth in the circuit so that it has the effect of flowing through the capacitor and often we say that ac flows through a capacitor, but actually there is no electron flow across the dielectric of a capacitor unless the capacitor breaks down.

(ag) Capacitive reactance.

(ah) $Xc = \frac{1}{6.28 \times f \times C}$ where f is in cycles/second and C is in farads. Also

$$Xc = \frac{159{,}000}{f \times C}$$

where f is in cycles/second and C is in microfarads.

(ai) To find the capacitive reactance of the capacitor we use the formula:

$$Xc = \frac{159{,}000}{f \times C}$$

and substituting 150 for f and .02 for C we get:

$$Xc = \frac{159{,}000}{150 \times .02}$$

$$Xc = \frac{159{,}000}{3}$$

$$Xc = 53{,}000 \text{ ohms}$$

(aj) .32 mfd. To find the capacity of capacitors which are connected in parallel you simply add the capacities.

(ak) To find the capacity of the two capacitors connected in series we use the formula:

$$C_t = \frac{C1 \times C2}{C1 + C2}$$

and substituting 27 pf for C1 and 56 pf for C2 we get:

$$C_t = \frac{27 \times 56}{27 + 56}$$

$$C_t = \frac{1512}{83}$$

$$= 18.2 \text{ pf}$$

(al) Ohms.

(am) The complete statement should be: Increasing the capacity or the frequency in the circuit will result in a lower capacitive reactance and decreasing the frequency or capacity will result in a higher capacitive reactance.

(an) The time-constant of an rc circuit is the length of time it takes the capacitor to charge up to approximately 63% of the total source voltage.

(ao) .044 seconds. To find the time-constant you multiply the resistance in megohms by the capacity in microfarads. 2.2 × .02 = .044 seconds.

(ap) 2 seconds. The time-constant will be exactly the same because the charging voltage has no effect on the time-constant. The time-constant is determined solely by the value of the resistance and the capacitor in the circuit. In each case the capacitor will charge up to 63% of the applied voltage in 2 seconds.

(aq) The current will lead the voltage by 90°.

(ar) The current will lead the voltage by some angle less than 90°.

(as) 15 volts. You can use either a vector solution to get this answer or the mathematical solution using the formula:

$$Eg = \sqrt{E_R{}^2 + E_C{}^2}$$

$$Eg = \sqrt{9^2 + 12^2}$$

$$Eg = \sqrt{81 + 144}$$

$$Eg = \sqrt{225}$$

$$= 15 \text{ volts.}$$

(at) By impedance we mean the

total opposition to the flow of ac in the circuit. The impedance will be the vector sum of the opposition offered by the resistance plus that offered by the capacitive reactance of the capacitor.

(au) We can find the impedance using the formula:

$$Z = \sqrt{R^2 + X_C^2}$$

and substituting 6 ohms for R and 8 ohms for X_C we have:

$$Z = \sqrt{6^2 + 8^2}$$

$$Z = \sqrt{36 + 64}$$

$$Z = \sqrt{100}$$

$$Z = 10 \text{ ohms.}$$

Lesson Questions

Be sure to number your Answer Sheet B107.

Place your Student Number on every Answer Sheet.

1. Name three types of solid dielectric capacitors.

2. Express .0001 mfd (microfarads) in pf (picofarads).

3. Can you use a .01-mfd, 600-volt capacitor in place of a .01-mfd, 400-volt capacitor?

4. What will happen to an electrolytic capacitor if you install it in a circuit with the wrong polarity?

5. What is the capacitive reactance of a 10-mfd capacitor at a frequency of 10 cycles?

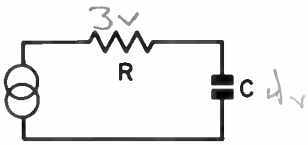
6. What is the total capacity of a 6-mfd capacitor and an 8-mfd capacitor connected in parallel?

7. What is the total capacity of two 8-mfd capacitors connected in series?

8. What do we mean by the time-constant of an RC circuit?

9. What is the phase relationship between the voltage across a capacitor and the current flowing through the capacitor in an ac circuit?

10. If the voltage across R is 3 volts and the voltage across C is 4 volts, what is the generator voltage in the circuit shown?



THE VALUE OF COURTESY

A recent survey showed that people complain more about discourteous clerks than about any other fault a business could have. In fact, many people pay extra at higher-priced stores just to get the courtesy and respect they feel entitled to.

Regardless of whether you work for someone else or have a radio business of your own, plain ordinary courtesy can bring many extra dollars to you.

Courtesy becomes a habit if practiced long enough. Be courteous to everyone -- to members of your family, to those who don't buy from you, even to the very lowest persons who serve you -- then you can be sure you'll be courteous when it really counts.

Give your courtesy with a smile. There is an old Chinese proverb which says, "A man who doesn't smile shouldn't keep a shop." And you're keeping a "shop" even if you are selling only your ability and knowledge to an employer. A friendly smile is itself courtesy of the highest type, bringing you unexpected returns in actual money as well as friendship.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-207

ACHIEVEMENT THROUGH ELECTRONICS

nri

HOW RESISTORS, COILS AND CAPACITORS ARE USED TOGETHER

B108

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

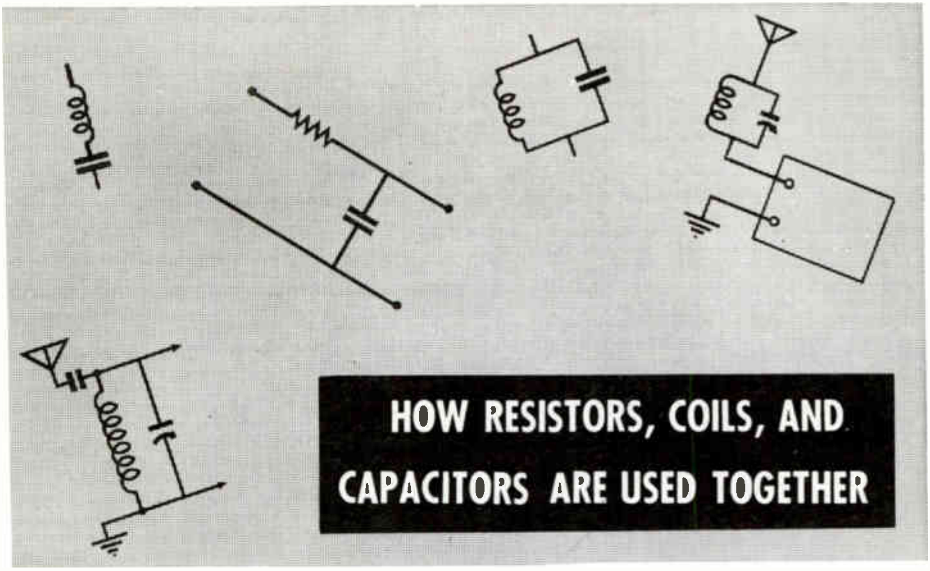
HOW RESISTORS, COILS, AND CAPACITORS ARE USED TOGETHER

B108

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

JDL3M1174 1974 EDITION Litho in U.S.A.



HOW RESISTORS, COILS, AND CAPACITORS ARE USED TOGETHER

In the preceding lessons you have studied resistors, coils and capacitors. Other than tubes and transistors, these are the three most important parts you will run into in electronic equipment. In your studies of these parts you have been primarily concerned with the part itself, with its characteristics and briefly with how it is made. In electronic equipment these parts are seldom used alone; in most cases two or more of these parts will be used together. For example, in coupling circuits, you will often use combinations of resistors and capacitors to feed the signal from one audio stage to a second audio stage. Circuits of this type are referred to as RC coupling circuits.

You will run into circuits where coils and resistors are used together. Circuits of this type are referred to as RL circuits. You will also encounter circuits where coils and capacitors are used together. These circuits are referred to as LC circuits. In this lesson we will study all three types of circuits but will spend most of the time on LC circuits because these are perhaps the most important of the three.

From your studies of the coil and the capacitor you might remember that in many ways one is the opposite of the other. For example, in a circuit having only inductance, you will recall that the voltage leads the current by 90°. On the other hand, in a circuit having only capacitance, the voltage lags the current by 90°. In other words, we will get the exact opposite effect; with a coil we have a 90° leading voltage and with a capacitor a 90° lagging voltage.

In circuits where coils and capacitors are used together the two more or less work against each other. In some of these circuits, the circuit may act like a coil and the voltage will lead the current by some angle less than 90°. In other circuits you may find that the circuit acts like a capacitor and the voltage lags the current by some value less than 90°. In still other circuits the effect of the coil will cancel the effect of the capacitor so that the voltage and the current will be in phase. Circuits of this type are called resonant circuits. Resonant circuits are extremely important. They are used in radio and TV receivers to separate the various stations. In other words, a resonant circuit is used in a broadcast-band receiver to select the station you want to listen to and at the same time reject the unwanted stations. Resonant circuits are used in TV receivers to tune the set to the channel you want. Without resonant circuits there would be no way to separate the stations.

There are two types of resonant circuits. One is called a series-resonant circuit, and the other is called a parallel-resonant circuit. Whether a circuit is series-resonant or parallel-resonant depends upon how the voltage is applied to the coil and the capacitor in the circuit. Both types are important; we will study both and you will soon see how to distinguish one type from the other. Since it is a little easier to see exactly what is happening in a series-resonant circuit than it is in a parallel-resonant circuit, we will study the series-resonant circuit first.

Series-Resonant Circuits

A resonant circuit is a circuit in which the inductive reactance of the coil is equal to the capacitive reactance of the capacitor. When the voltage is applied to the coil and the capacitor in series we call the circuit a series-resonant circuit. To help you get a clear understanding of what a series-resonant circuit is, let us start with a simple series circuit and review some of the things you already know.

A SERIES CIRCUIT

Let's begin with the circuit shown in Fig. 1, consisting of a 500-cycle ac generator that is generating a voltage of 120 volts. Across this generator we will connect a variable resistor, which is set so that it has a resistance of 120 ohms. If we have a voltmeter connected across the resistor, the voltage reading will be close to 120 volts. This is true because the resistor is connected directly across the generator, and the voltage being supplied by the generator is 120 volts. An ammeter connected in series with the resistor will indicate that the current flowing in the circuit is 1 ampere. Actually, we do not need the ammeter to tell us this, because we know that we can determine the current by dividing the voltage by the resistance. Remember this formula for Ohm's Law:

$$I = \frac{E}{R}$$

Therefore

$$I = \frac{120}{120} = 1 \text{ amp}$$

In this simple circuit we know that the current will be in phase with the voltage, because the only element connected in the circuit is resistance. This means that the voltage across the resistor will reach its peak at the same time that the current flowing in the circuit is at a maximum.

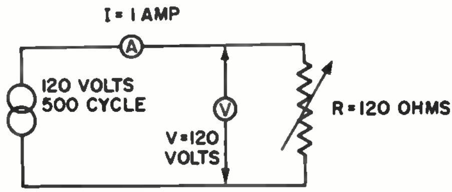


Fig. 1. A simple series circuit.

Remember that the voltmeter and the ammeter are measuring the effective value (or the rms value) of the ac voltage and current and that during part of the cycle, the generator voltage actually exceeds 120 volts. You will recall that the peak voltage of an ac sine wave will be 1.41 times the effective voltage. Therefore, even though the meter is reading 120 volts, the voltage will be greater than this value during part of the cycle and less than this value during the other part of the cycle. The voltage will actually be zero twice each cycle. Likewise, the current is the effective current and its actual value will be greater than 1 amp twice each cycle and less than 1 amp the remainder of the time. The current will reach a peak value of 1.41 amps twice each cycle and will actually drop to 0 twice each cycle.

AN RL CIRCUIT

Now let us see what happens when we modify the simple circuit of Fig. 1 by adding a 100-mh (millihenry) coil in series with the resistor as shown in Fig. 2. In Fig. 1 we showed the resistor as a variable resistor. We did this deliberately because you know that any coil will have some resistance. If the 100-millihenry coil has a resistance of 10 ohms, we can adjust the variable resistor so that its resistance is only 110 ohms. Now we will still have a total of 120 ohms resistance in the circuit: the 110 ohms from the resistor and the 10 ohms from the coil. In Fig. 2, we have represented the total resistance made up of the coil resistance and the resistance of the variable resistor as only one 120-ohm resistor. Notice that in Fig. 2 we have the ammeter in series with both the coil and resistor and have added a second voltmeter across the coil. But look at the readings on the meters! The ammeter shows that the current flowing is .35 amp. The voltmeters show that the voltage across the resistor is 43 volts, and the voltage across the coil is 112 volts. You have already seen this type of circuit, and probably know why we have obtained these readings, but let's go through this again to be sure you understand why such

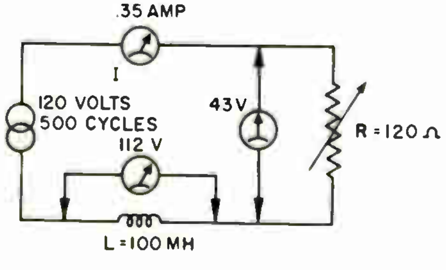


Fig. 2. A simple RL series circuit.

voltage readings are obtained. You may at first think this looks complicated, but actually it is not nearly as difficult as it might appear.

First, we have a series circuit. We know that the total opposition to the current flowing in the circuit will be the impedance of the circuit, which in turn is made up of the 120-ohm resistance plus the inductive reactance of the coil. We can find the inductive reactance of the coil from the formula:

$$X_L = 6.28 \times F \times L$$

Since 100-mh is .1 henry, L = .1 and F = 500 cycles, we have:

$$X_L = 6.28 \times 500 \times .1$$
$$= 314 \text{ ohms}$$

As you might expect, it is extremely difficult to make a coil with an inductance of exactly 100 millihenrys. Furthermore, we are not interested in exact calculations in most electronic circuits. To simplify things, we are going to call the inductive reactance of the coil 315 ohms. The impedance of the circuit is therefore equal to 120 ohms plus the reactance of 315 ohms. To show this by means of symbols, electronics men write this as

$$Z = 120 + j315$$

The letter j is used to indicate that 315 ohms is a reactive component which cannot be added directly to 120 ohms to get the total impedance of the circuit. If we simply wrote the impedance as 120 + 315, it would be very easy to forget that 120 was resistance and 315 was reactance and simply add the two and get 435 ohms, which of course, is incorrect.

The value of Z can be determined either by vectors as shown in Fig. 3 or from the formula

$$Z = \sqrt{R^2 + X^2}$$

In this case R is 120 ohms, and X is the inductive reactance, which is 315 ohms. So we have:

$$Z = \sqrt{120^2 + 315^2}$$

Remember that the small 2 to the right of 120 and 315 means that the number is to be squared, and that 120 squared is equal to 120 times 120. If you want to go through the arithmetic, the actual operations of squaring these numbers are as follows

$$\begin{array}{r} 120 \\ \times\ 120 \\ \hline 2400 \\ 120 \\ \hline 14400 \end{array} \qquad \begin{array}{r} 315 \\ \times\ 315 \\ \hline 1575 \\ 315 \\ 945 \\ \hline 99225 \end{array}$$

Now that we have the value of 120 and 315 squared, we have

$$Z = \sqrt{14{,}400 + 99{,}225}$$

and by adding these we will get

$$\begin{array}{r} 14400 \\ +\ 99225 \\ \hline 113625 \end{array}$$

Therefore $Z = \sqrt{113{,}625}$

To get the value of Z we have to take the square root of 113,625. This is done by first writing the number down and then marking it off in groups of two numbers working from the decimal point to the left. The steps in getting this square root are as follows:

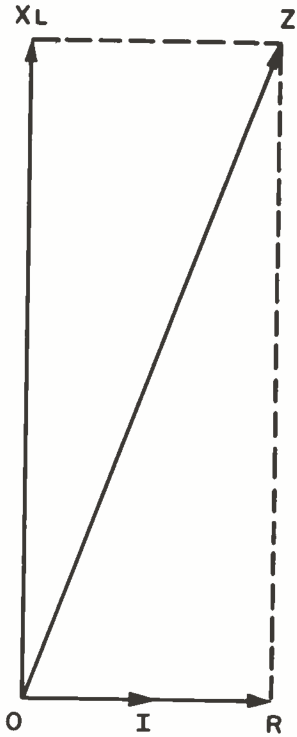


Fig. 3. Vector addition of R and X_L using a scale of 100 ohms = 1 inch, when R = 120 ohms, and X_L = 315 ohms. R is drawn 1.2 inches long, and X_L is drawn slightly less than 3.2 inches long. The impedance vector Z is then found by completing the rectangle and drawing Z from zero to the junction of the dotted lines. Z will measure between 3.25 and 3.5 inches, giving a Z of about 340 ohms.

$$\begin{array}{rr} & \sqrt{11'\,36'\,25'} \quad \underline{|337} \\ & 9 \\ \underline{63|} & \overline{2\,36} \\ & 1\,89 \\ \underline{667|} & \overline{47\,25} \\ & 46\,69 \\ \hline & 56 \end{array}$$

Therefore, the impedance of the circuit is 337 ohms. We need not be concerned if there is a slight remainder when we work out the square root because electronic parts have considerable tolerance, and 337 ohms is close enough. As a matter of fact, in many practical problems you could probably round this off either to 335 or 340 ohms. However, we'll use 337 ohms.

Now that we know the total impedance in the circuit, we can determine what the current flowing in the circuit would be in this manner:

$$I = \frac{E}{Z} = \frac{120}{337} = .356 \text{ amp}$$

The ammeter we have connected in the circuit should and does read about .35 amp.

Now that we know the current flowing in the circuit, we can determine the voltage that will appear across the coil by multiplying the reactance of the coil by the current flowing in the circuit. In other words:

$$E_L = I \times X_L$$

Multiplying .356 × 315 we get 112.14 volts. We do not have to be concerned with such accuracy as this; we will simply call the voltage 112 volts. As a matter of fact this is the voltage indicated by the voltmeter connected across the coil; you cannot get a more accurate reading. Similarly, the voltage across the resistor is

$$E_R = I \times R$$

which is equal to .356 × 120, or 42.72 volts. We will round this off to 43 volts, which is what the meter will indicate.

Each of the steps that we have shown in the preceding example is important. It is not particularly necessary that you sit down and follow through the multiplication and division unless you want to do so. If you expect to go into radio and TV servicing you will have no occasion to do this type of work, but if you intend to go into industry as an electronics technician, you should be sure you understand the various steps in this circuit explanation.

Whether you go through the mathematics or not, there are several important points you should see. The current flowing in the circuit is limited by the impedance of the circuit. The impedance, as you know, is the total opposition to current flow in the circuit. The voltage that will appear across each part in the circuit will depend upon the resistance or reactance of that part and upon the current flowing in the circuit. Also, you should remember that in this example, although it may appear that the sum of the voltage across the coil plus the voltage across the resistance is greater than the source voltage, the meters are measuring the effective value of the voltage and that these voltages are not in phase. These voltages are ac voltages and hence continually changing, but the voltage across the resistor plus the voltage across the coil at any given instant is exactly equal to the source voltage at that instant.

THE RC CIRCUIT

Now, let's remove the 100-mh coil from the circuit and put a 1-mfd capacitor in its place. We will readjust the variable resistance so that the total resistance of the circuit is 120 ohms. We will then have the circuit shown in Fig. 4. Notice that

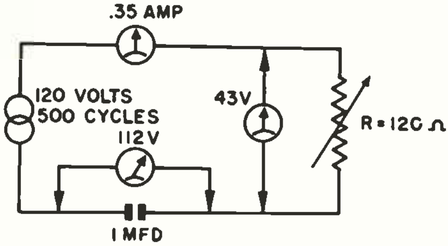


Fig. 4. A simple RC series circuit.

the voltage across the resistance is the same as in Fig. 2, and also notice that the voltage across the capacitor in Fig. 4 is the same as the voltage across the coil in Fig. 2.

We can verify all the meter readings shown in Fig. 4 as we did for the circuit shown in Fig. 2. To do this, we start by finding the reactance of the 1-mfd capacitor at 500 cycles from the formula:

$$X_C = \frac{1}{6.28 \times F \times C}$$

The capacitive reactance of the capacitor turns out to be 318 ohms. We will round this off to 315 ohms, because it is unlikely that the capacitor will have a capacity of exactly 1 microfarad. You do not have to go through the solution of this formula unless you want to, but for the benefit of those who want to work it out step by step we'll go through it.

First, remember that in the formula the frequency must be in cycles and the capacity must be in farads. Therefore, F = 500 cycles and C = 1 mfd = .000001 farad. Substituting these values in the formula we have

$$X_C = \frac{1}{6.28 \times 500 \times .000001}$$

Now we can start by multiplying 6.28 × 500 and we will get

$$\begin{array}{r} 6.28 \\ \times\ 500 \\ \hline 3140.00 \end{array}$$

Next, multiplying 3140 by .000001, we will get

$$\begin{array}{r} 3140 \\ \times\ .000001 \\ \hline .003140 \end{array}$$

The next step in our problem is to divide 1 by .003140. To perform the division we must get rid of the decimal point; to do this we move the decimal point six places in the divisor and six places in the dividend so that our problem becomes 1000000 divided by 3140:

$$.003140 \overline{)1.000000}$$

$$\begin{array}{r} 318 \\ 3140 \overline{)1000000} \\ \underline{9420} \\ 5800 \\ \underline{3140} \\ 26600 \\ \underline{25120} \\ 1480 \end{array}$$

As we mentioned we will round off the value of the capacitive reactance to 315 ohms. The difference is so small that whether it is 315 or 318 you won't be able to detect any difference in the meter readings in Fig. 4. Using 315 ohms as the capacitive reactance makes the total impedance of the circuit

$$Z = 120 - j315$$

Notice that in this case we have used -j to indicate that 315 ohms is a capacitive reactance. Actually, whether the sign is minus or plus will make no difference in determining the value of the impedance. We still use the formula:

$$Z = \sqrt{R^2 + X^2}$$

In this case, X is the capacitive reactance. Even though it is shown as -j315, it will still be +when squared, because any number squared is positive. Our formula is therefore the same as before:

$$Z = \sqrt{120^2 + 315^2}$$

The impedance works out again to be 337 ohms. The current flowing in the circuit can then be found by dividing the voltage by the impedance; this is 120 ÷ 337, or .356 amp. We can find the voltage across the resistor by multiplying the current, .356 amp, by the resistance, 120 ohms. Thus, .356 × 120 = 43 volts. Similarly, the voltage across the capacitor is .356 × 315, or 112 volts.

It is worthwhile to notice that since the circuit shown in Fig. 2 contains inductance and resistance the current will be lagging the voltage. In the circuit shown in Fig. 4 the current will be leading the voltage since we have capacitance and resistance. Nevertheless, the total impedance is the same in the two circuits; therefore, the currents flowing in the circuits are equal, which accounts for the equal voltage appearing across the resistors in each circuit.

Now let us see what happens when we go one step further and put the 100-mh coil in series with the 120-ohm resistor and the 1-mfd capacitor.

THE RESONANT CIRCUIT

In the circuit shown in Fig. 5, we have 120-ohm resistance in series with the 100-mh coil and the 1-mfd capacitor, and the series combination connected across the 120-volt, 500-cycle generator. The inductive reactance of the coil and the capacitive reactance of the capacitor will each be approximately 315 ohms as before, because the frequency in the circuit has not changed. Therefore, the total impedance in the circuit will be

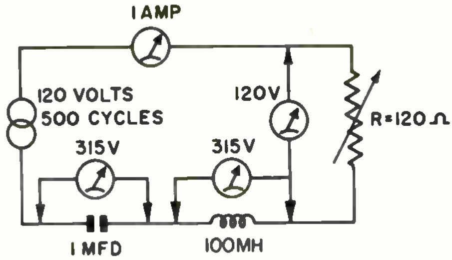


Fig. 5. A series-resonant circuit.

$$Z = 120 + j315 - j315$$

We have already pointed out that capacitive reactance is essentially the opposite of inductive reactance. We have indicated this by using +j to represent inductive reactance and -j to represent capacitive reactance.

In the expression for the impedance of the circuit you see +j315 and -j315, and, as you might expect, these two cancel so that the total impedance of the circuit is equal to the resistance of the resistor alone, or 120 ohms.

Now let's see what happens to the voltages and current throughout the circuit. First, look at the current flowing; it is 1 amp. It is higher than it was in Fig. 2 and Fig. 4. The reason for this is that the current is equal to

$$I = \frac{E}{Z}$$

Z in Fig. 5 is 120 + j315 - j315, or 120 ohms. Therefore I = 120 ÷ 120, or 1 amp.

The voltage across the coil is 315 volts. This is equal to the current times the inductive reactance: 1 amp × 315 ohms = 315 volts. Similarly, the voltage across the capacitor will be equal to 315 volts. The voltage across the resistor is 120 volts. Notice what we now have across the capacitator and across the inductance: a voltage several times the source voltage. This is referred to as a "resonant voltage step-up." In other words, the voltage across the coil and the voltage across the capacitor in a series-resonant circuit may be several times the source voltage.

In the preceding example we saw that the impedance of the circuit at resonance is equal to the resistance in the circuit. Now you might wonder what happens when you change the resistance in the circuit. If you reduce the value of the resistance in the circuit, you will reduce the impedance. If we cut the resistance in half so that the total resistance in the circuit is only 60 ohms, the current flowing in the circuit will be doubled. We will then have the situation shown in Fig. 6. Here the

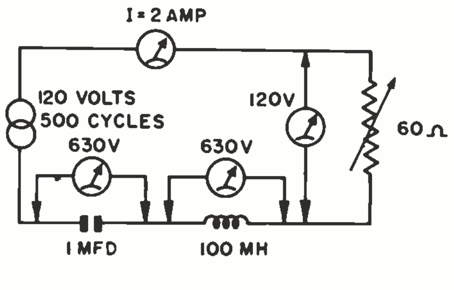


Fig. 6. Reducing the resistance in the series-resonant circuit results in a greater current flow and a higher resonant voltage step-up.

voltage drop across the capacitor is equal to 2 amps × 315 ohms, or 630 volts. Similarly, the voltage appearing across the coil will be 630 volts and the voltage across the resistor will be 2 amps × 60 ohms, which is equal to 120 volts as before. Now we have an even greater resonant voltage step-up than we had previously. Reducing the resistance still further will result in an even greater voltage appearing across the coil and across the capacitor.

In a practical resonant circuit, we will have only a capacitor and a coil in series. However, there will always be resistance in the circuit because the coil is made by winding turns of copper wire on a coil form, and the copper wire has resistance. However, the lower that resistance can be made, the lower the impedance of the circuit and the greater the resonant voltage step-up will be.

This can be expressed in another way. The Q of a coil is equal to the inductive reactance divided by the resistance of the coil. In other words:

$$Q = \frac{X_L}{R}$$

In a coil where the resistance is low, the value of Q will be high. Since the lower the resistance we have, the greater the resonant step-up voltage will be, we say that there will be more resonant voltage step-up in a high-Q circuit than in a low-Q circuit.

Series-Resonant Facts.

Before going ahead with our study of series-resonant circuits, we should review several of the things that we have already discussed.

First, in a series-resonant circuit at resonance the inductive reactance is exactly equal to and cancels the capacitive reactance in the circuit so that the impedance of the circuit is equal to the resistance in the circuit. The impedance in the circuit will therefore be low, and the current flowing in the circuit will be high. Second, there is a resonant voltage step-up in a series-resonant circuit.

Remember these two facts -- they are important: in a series-resonant circuit you will have low impedance and high current and the voltage appearing across either the capacitor or inductance may be several times the source voltage.

Resonance occurs when the inductive reactance of the coil exactly equals the capacitive reactance of the capacitor. In other words,

$$X_L = X_C$$

$$6.28 \times F \times L = \frac{1}{6.28 \times F \times C}$$

and from this we can get:

$$F = \frac{1}{6.28 \times \sqrt{L \times C}}$$

This is the frequency at which a coil and a capacitor will be resonant.

VOLTAGE AND CURRENT WAVEFORMS IN THE SERIES-RESONANT CIRCUIT

In any series circuit, the current is always the same in all parts of the circuit. This is true whether it is a resonant circuit or a simple series circuit consisting of a number of resistors. Therefore, in a circuit such as the one shown in Fig. 7A the current flowing through the generator will be equal to the current flowing through the coil at all times, whether the circuit is resonant or not. This current is also equal to the current flowing through the resistor and to the current flowing through the capacitor.

If the circuit shown in Fig. 7A is a resonant circuit, we can make use of the fact that the current is the same at all times in all parts of the circuit to get a better idea of what is happening. We can do this by studying the voltage and current waveforms throughout the circuit. For example, we have shown two cycles and have identified a number of points on these cycles by numbers in Fig. 7B. This waveform represents the entire current flowing in the series-resonant circuit.

Look at the current waveform at point 1 and notice that the current begins swinging in a positive direction. Let's assume that at this instant the voltage at terminal a of

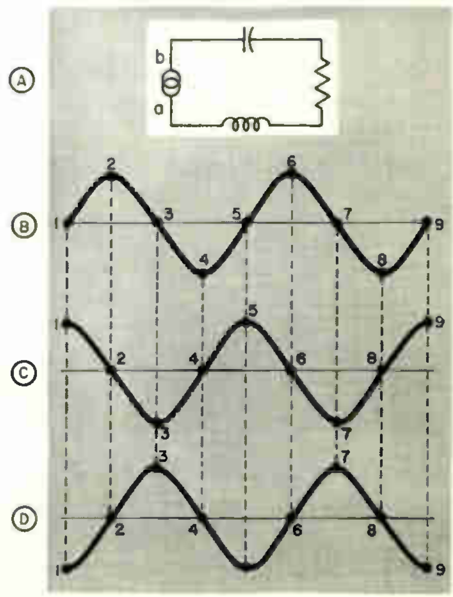


Fig. 7. Series-resonant circuit waveforms.

the generator begins to swing in a negative direction so that current starts flowing at terminal a through the coil and around the circuit back to terminal b of the generator. As the waveform increases in amplitude from point 1 to point 2, current increases in the circuit going from 0 at point 1 to its maximum value at point 2.

Although we have started the waveform shown in 7B at point 1, we are actually representing what happens in a series-resonant circuit in which the current has been flowing for some time. In other words, the circuit has been connected to the generator (which is producing the current) and we simply started to analyze what is happening in the circuit at the given instant designated at point 1 on the curve B. At this instant the current flowing in the circuit has just reached 0 and has begun to increase towards point 2 as electrons move from terminal a of the generator around the circuit towards terminal b of the generator.

Fig. 7B is a sine wave. A sine wave changes value at a maximum rate at the instant it is going through 0. In other words, at point 1 the rate at which the value of the current (represented by the sine wave) changes is at a maximum. As the actual current amplitude increases towards point 2, the rate of change decreases until for an instant at point 2 there is no change in current. The current has reached its maximum value, it remains constant for just an instant at point 2, and then begins to decrease. The rate at which it decreases increases until it is changing at its maximum rate when it reaches point 3. For just an instant the current drops to 0 at point 3. Thus, the current goes through a complete change - the instant before it reaches point 3 it is flowing in one direction, it drops to 0 exactly at point 3 and then an instant later it is flowing in the opposite direction.

Now let us consider what effect this changing current has on the voltage across the coil. Remember that the self-induced voltage in a coil depends upon the rate at which the current flowing through the coil changes. Therefore, at point 1, since the current is changing at its maximum rate, the self-induced voltage induced in the coil will be at maximum. As the current increases in amplitude towards point 2, the rate at which it is changing decreases so the voltage induced in the coil decreases. At point 2, where the current is not changing at all, the voltage induced in the coil will be 0. This is shown by the portion of curve C between points 1 and 2. Now, as the current begins to decrease from point 2 to point 3 the rate at which the current is changing increases. Since the current is decreasing, a voltage is induced in the coil which tends to oppose the direction in which the current is changing. Therefore, as the current goes from point 2 to point 3 (as shown in the curve at B), the voltage across the coil will increase from 0 at point 2 to its maximum value at point 3 (as shown in the curve at C).

Notice that the voltage across the coil is leading the current by 90°. If you remember your earlier studies, this is exactly what you might expect. As the current goes through the remainder of its cycle (as shown from point 3 on over to point 9), the voltage across the coil will at all times be 90° ahead of this current.

Now let us consider what must be happening across the resistor and capacitor in the circuit. At point 1 on curve B the actual value of the current flowing is 0. Therefore, the voltage across the resistor must be 0 because it will be in phase with the current. We have not drawn a wave shape to represent the voltage across the resistor because it will be in phase with the current waveform shown at B at all times.

Since the voltage across the generator is 0 at point 1 and the voltage across the resistor is also 0 at point 1, we immediately see that the voltage across the capacitor must be exactly equal to and opposite to the voltage across the coil. We know this must be true because in any closed circuit of this type the algebraic sum of the voltage drops around the circuit must be equal to 0. Indeed, if the circuit has been in operation the capacitor would be charged with a polarity equal to and opposite to the voltage across the coil at the instant that the current was at 0, as at point 1. As the current increases from point 1 to point 2, electrons will flow into the capacitor to reduce the charge until at point 2 the charge on the capacitor will be exactly equal to 0. As current continues to flow in the same direction (as shown from point 2 to point 3 on curve B) electrons will continue to flow into one plate of the capacitor and charge it (as shown between points 2 and 3 on curve D).

Notice what has happened. At each instant, the voltage across the capacitor is equal to and opposite to the voltage across the coil. Also notice that the voltage across the capacitor is lagging the current flowing in the circuit by 90°. This is as we should expect; when you studied capacitors in earlier lessons, you learned that the voltage lags the current in a capacitor by 90°.

The curves in Fig. 7 show what happens in a series-resonant circuit and why the voltages around the circuit (if they are measured separately, then added together) will be equal to a value greater than the source voltage. If you measure the capacitor voltage you'll be measuring the rms value of the waveform shown at D. If you measure the coil voltage you will be measuring the rms value of the curve shown at C. While these voltages do exist across the components, they are always 180° out of phase and therefore as far as the total voltage across the two is concerned they cancel each other.

In a series-resonant circuit that has a pure inductance and a pure capacitance in series with a resistor, the voltage across the inductance is exactly cancelled by the voltage across the capacitance. In this case, the applied voltage is developed across the resistor. However, pure inductors and capacitors do not exist. All coils have some resistance and all capacitors have some leakage. Therefore, the voltages across the coil and capacitor do not exactly cancel. Thus, a small voltage can be measured across the coil-capacitor combination.

Before going on to the next section of the lesson it would be worthwhile to study carefully the waveforms shown in Fig. 7 to be sure you understand what is happening. It might make it easier for you if you can redraw the circuit and place a resistor next to the generator instead of between the coil and the capacitor. Actually, insofar as the current flow in the circuit and the voltage across the individual parts in the circuit are concerned, it makes no difference where the parts are placed. However, to see how the voltages across the coil and capacitor are cancelling, it might be easier to see this if you move the resistor so that it is not between the two.

VARYING L, C, F, AND R

When the inductive reactance in a series circuit is equal to the capacitive reactance, the circuit is at resonance, the two cancel each other, and we will have a low-impedance circuit in which we will get a high current flow. Let us see what happens if we use other values of C and L and if we vary the frequency of the voltage applied to the circuit. Let's start by seeing what will happen when we vary the value of the capacitor.

Varying C.

We can vary the value of C by inserting capacitors of different sizes in the circuit in place of the 1-mfd capacitor, while leaving the source frequency at 500 cycles and the coil inductance at 100 millihenrys. The current in the circuit will vary as shown in Fig. 8. Notice that we have the highest current flowing exactly at resonance. As the capacity is reduced, the current drops off rather sharply. It is important for you to realize that as the capacity is made smaller than it should be for resonance, the capacitive reactance in the circuit increases so that it is greater than the inductive reactance. Therefore, the inductive reactance does not completely cancel out all the capacitive reactance in the circuit, and the resonant circuit begins to act like a circuit having only capacity. In other words, the current flowing in the circuit will lead the voltage.

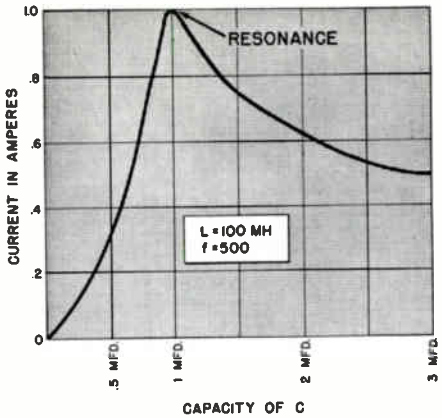


Fig. 8. How current varies in a series-resonant circuit when L is 100 mh, F is 500 cycles, and C is varied from 0 to 3 mfd.

As the capacity is increased above 1-mfd, the capacitive reactance decreases so that the inductive reactance in the circuit is greater than the capacitive reactance. Therefore, the inductive reactance cancels out all of the capacitive reactance and there is still some inductive reactance left over. The circuit begins to act like a circuit having only inductance in it and the current will lag the voltage.

In either case, when the capacity is too small or too high for resonance, the impedance of the circuit is greater than it is at resonance; this accounts for the reduction in current flow in the circuit.

Varying L.

If the inductance in the circuit is varied instead of the capacity, the current will vary as shown in Fig. 9. Here, when the size of the inductor or coil is reduced, the inductive reactance is decreased, and it will not completely cancel out the capacitive reactance in the circuit. The circuit will act like a capacitor, and the current will lead the voltage.

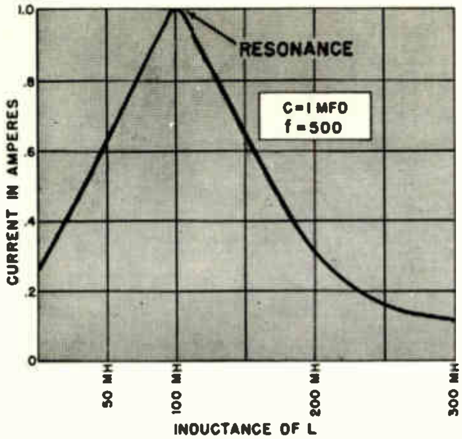


Fig. 9. How current varies in a series-resonant circuit when C = 1 mfd, F = 500 cycles, and L is varied from 0 to 300 mh.

When the inductance is made greater than 100-mh, then the inductive reactance is greater than the capacitive reactance and the circuit will act as an inductance. This means that the current will lag the voltage. Again, the current will be lower than at resonance when the inductance is made either too large or too small because the impedance in the circuit increases. Lowest impedance is obtained in a series-resonant circuit at resonance, when the inductive reactance cancels the capacitive reactance. Current will always be at maximum at this point.

Varying Both L and C.

The combination of a 100-mh coil and a 1-mfd capacitor is not the only combination that will give resonance at 500 cycles. If we double the inductance of the coil by using a 200-mh coil in the circuit, the inductive reactance of the coil will be doubled. It will be twice 315 ohms, or 630 ohms. If we reduce the size of the capacitor from 1-mfd to .5-mfd, we will also double the reactance of the capacitor so that its reactance will now be 630 ohms at 500 cycles. This means that with a 200-mh coil and a .5-mfd capacity, we will again have a situation where the inductive reactance will be equal to the capacitive reactance; therefore, these components will be resonant at 500 cycles.

You can see that many values of coil and capacitor may be used to obtain resonance at 500 cycles. Once we select a coil, it will have a certain inductive reactance at a frequency of 500 cycles. All we need is to obtain a capacitor that will have a capacitive reactance equal to the inductive reactance of the coil at this frequency, and we will have a resonant circuit at 500 cycles.

Varying F.

If we go back to the 100-mh coil and the 1-mfd capacitor and then vary the frequency of the voltage applied to this combination, we will obtain a curve like the one shown in Fig. 10. Notice that in this case we have maximum current flow at resonance, because at resonance the impedance of the circuit is lowest. When the frequency applied to the coil and capacitor combination is less than 500 cycles, the capacitive reactance of the capacitor is greater than the inductive reactance of the coil, so the combination acts like a capacitor and the current will lead the voltage.

When the frequency applied to the combination is above 500 cycles, then the inductive reactance of the coil is greater than the capacitive reactance of the capacitor, and the circuit will act like a coil. In either case, the impedance of the circuit will be minimal at resonance and higher either above or below the resonant frequency. The fact that the impedance reaches its minimum at the resonant frequency of 500 cycles is the reason the current is at maximum.

Varying R.

Under certain circumstances, the total resistance in the circuit may be varied. As the resistance in the circuit is changed, the Q of the circuit changes and the current that will flow in this circuit changes. In Fig. 11, we have shown three resonance curves. The curve marked A is the one we had in Fig. 10 with a 100-mh coil, a 1-mfd capacitor, and a series resistance of 120 ohms. In the curve marked B, the inductance and capacity are the same, but the series resistance is 100 ohms.

Changing the resistance has an important effect on the shape of a resonant curve, as shown in Fig. 11.

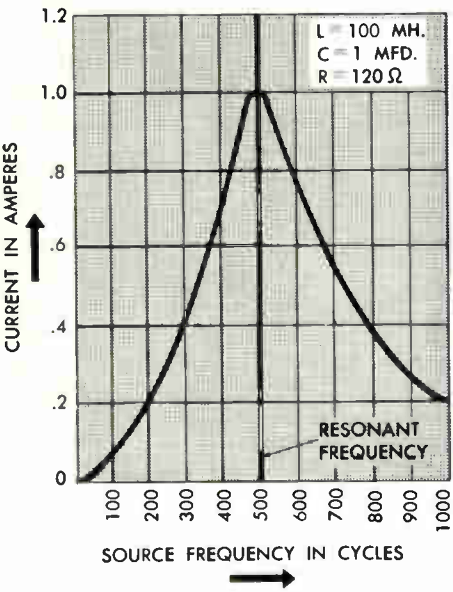


Fig. 10. How current varies in a series-resonant circuit when C = 1 mfd, L = 100 mh, R = 120 ohms, and F is varied from 0 to 1000 cycles.

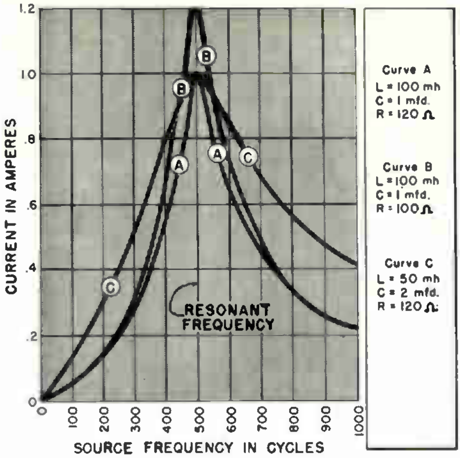


Fig. 11. How the current varies in a series-resonant circuit when F is varied, and (A) C = 1 mfd, L = 100 mh, and R = 120 ohms; (B) C = 1 mfd, L = 100 mh, and R = 100 ohms; (C) C = 2 mfd, L = 50 mh, and R = 120 ohms. In each case the product of L and C is the same, and resonance occurs at 500 cycles.

With a high-Q circuit, which is obtained when the resistance of the circuit is low, we have a very sharp curve. This means that we will get a much higher current flow at the resonant frequency of the circuit than we will get at a frequency either slightly above or slightly below the resonant frequency. However, with a low-Q circuit, we have a broad curve. This means that we can vary the frequency quite a bit either above or below the resonant frequency without causing a very great change in the current that will flow in the circuit.

The effect of resistance on a series-resonant circuit is important. Sometimes a series-resonant circuit may be used to select one frequency and reject all others. Obviously, if we have a high-Q circuit it will do a much better job of selecting one frequency and rejecting others than a low-Q circuit will. However, there are other instances when we are interested in selecting a band of frequencies rather than one particular frequency. In this case, a low-Q series-resonant circuit is used rather than a high-Q circuit which might not select the entire band or group of frequencies in which we are interested.

The curve marked C in Fig. 11 is the one we would obtain with a 50-mh coil, a 2-mfd capacitor and a series resistance of 120 ohms. Compare this curve with curve A. Notice that curve A is considerably sharper than curve C. The ratio of the inductance to the capacity is called the L to C ratio. In a series-resonant circuit a high L to C ratio will give you a sharper resonance curve than a low L to C ratio.

SELF-TEST QUESTIONS

(a) What is a resonant circuit?

(b) What do we mean when we say that the current flowing in the circuit in Fig. 1 is 1 amp?

(c) In a circuit such as the one shown in Fig. 2, why can we not simply add the voltage across the coil and the voltage across the resistor to find the total circuit voltage?

(d) The impedance of a circuit is given as $Z = 50 + j50$; what does the j mean?

(e) In an RC circuit, which part will have the greater voltage across it?

(f) In a series-resonant circuit, how will the phase of the voltage across the various components compare with the phase of the current flowing in the circuit?

(g) What do we mean by the Q of a coil?

(h) When a coil and a capacitor are connected in series, and the frequency of the voltage applied to them is varied, at what point will the current flowing in the circuit reach its maximum value?

(i) What effect on the current flowing in a series-resonant circuit will reducing the resistance in a circuit have?

Parallel-Resonant Circuits

When the source voltage for a resonant circuit is supplied across the coil and capacitor (that is, in parallel with them), as in Fig. 12, we have a

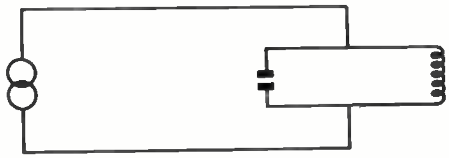


Fig. 12. A parallel-resonant circuit.

parallel-resonant circuit. In a parallel-resonant circuit, as in a series-resonant circuit, the inductive reactance of the coil is exactly equal to and cancels the capacitive reactance of the capacitor. This is essentially where the similarity between the two types of resonant circuits ends. In most respects, a parallel-resonant circuit acts in the opposite way to a series-resonant circuit. Let's investigate the characteristics of this type of resonant circuit and see why it performs as it does.

CIRCUIT CURRENT AND IMPEDANCE

In Fig. 13 we have shown a parallel-resonant circuit connected in series with a 120-ohm resistor across a 120-volt, 500-cycle generator. The coil has an inductance of 100 mh and the capacitor has a capacity of 1 mfd. All the conditions are as they were when we studied the series-resonant circuit. We have connected an ammeter in series with the resistor and the parallel-resonant circuit, a voltmeter across the coil and capacitor combination, and another voltmeter across the resistor.

In this type of circuit you can notice immediately that the current being supplied by the generator is very low. In addition, the voltage across the resistor is very low and nearly the entire source voltage appears across the resonant circuit.

The fact that the current flowing in the circuit is low immediately points out one important fact -- if the current flowing in the circuit is low, the impedance of the circuit must be high. In fact, this is the case; one of the most important characteristics of a parallel-resonant circuit is that at the resonant frequency it acts as a high-value resistance. Notice that this is just the opposite of a series-resonant circuit: at resonance a series-resonant circuit acts as a low resistance.

The fact that the parallel-resonant circuit acts like a high resistance explains why most of the source voltage appears across the resonant circuit and very little voltage appears across the 120-ohm resistor. The

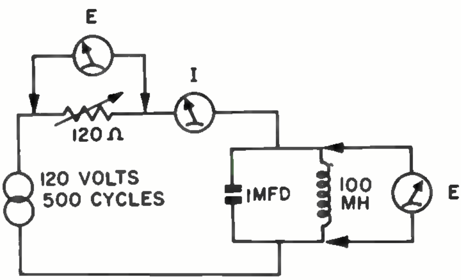


Fig. 13. In a parallel-resonant circuit, the current supplied by the generator is low, and most of the generator voltage appears across the resonant circuit.

120 volts supplied by the generator is simply divided between the resistor and the resonant circuit, with most of the voltage appearing across the higher resistance. Now let's study the parallel-resonant circuit in detail to see why it acts like a high value resistance.

Since the coil and capacitor in the parallel-resonant circuit are connected in parallel, there are two paths or branches through which current can flow. We call the path with the capacitor in it the capacitive branch, and the one with the inductance in it the inductive branch. If we connect an ac ammeter in each branch of the parallel-resonant circuit as shown in Fig. 14, we will discover that although we have a very low current being supplied by the generator, we have a very high current in each branch of the resonant circuit. You might wonder how this could be, but if we consider each branch separately we can see what is happening.

First, in a capacitive circuit we know that the current leads the voltage by 90°; in an inductive circuit the current lags the voltage by 90°. Therefore, the current flowing in the capacitive branch will be 180° out of phase with the current flowing in

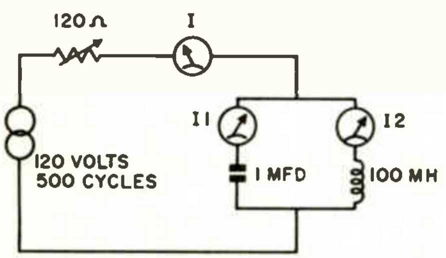


Fig. 14. Although the current supplied by the generator is low, the current flowing in each branch of a parallel-resonant circuit is high.

the inductive branch. The fact that the currents I_1 and I_2 are 180° out of phase means that the capacitor is discharging during one half cycle while the coil stores up electrical energy. During the next half cycle when the coil is releasing the electrical energy it has stored, the capacitor is charging. When the reactance of the coil is equal to that of the capacitor, as it will be at resonance, the energy stored by the capacitor equals the energy released by the coil; and during the other half of the cycle, the energy stored by the coil equals the energy released by the capacitor.

Thus the coil and the capacitor pass current back and forth to each other inside the resonant circuit. The actual amplitude of the current will depend upon the amount of resistance in the circuit. You know that the coil will have some resistance, and in addition, the leads connecting the coil and capacitor together have some resistance. However, because the resistance is usually kept quite low there can be a very high current flowing back and forth between the coil and the capacitor.

The fact that there is resistance in the circuit means that there will be some energy lost during each cycle. The very low current being supplied by the generator actually replaces the energy lost as heat because of the resistance in the resonant circuit.

The situation in the parallel-resonant circuit may be compared to the pendulum of a clock. The pendulum swings back and forth, and the current does essentially the same thing; it flows out of the coil into the capacitor and then back from the capacitor into the coil. If the pendulum in a clock were swinging freely, it would lose energy each oscillation due to friction and each arc would be smaller than the previous one; finally, it would come to rest. In a parallel-resonant circuit the current will flow back and forth and get smaller each cycle and eventually drop to zero unless some outside energy is supplied to it. The mechanical drive in the clock supplies the energy to the pendulum to keep it swinging; in the resonant circuit, the generator across the circuit supplies the energy to make up the losses in the circuit. Once the action of the current flowing back and forth in a parallel-resonant circuit has started, it will continue for a number of cycles until all the energy is used up in the resistance in the circuit. Similiarly, the pendulum of a clock will swing back and forth for a number of cycles once it is started in motion even if no additional energy is supplied to it.

The situation we have found in the parallel-resonant circuit exists at all times when a coil and capacitor are connected in parallel in an ac circuit. One feeds energy or current back into the circuit while the other draws current. Therefore, the current supplied by the generator at any instant will be the difference between the two currents. When the reactances are equal, as they are at resonance, then this current becomes the minimum current needed to make up the losses in the parallel-resonant circuit. Because the current does drop to a minimum value at resonance, the parallel-resonant circuit acts like a resistor of high ohmic value and reduces the line current supplied by the generator to a very low value.

At resonance, the inductive and capacitive currents are equal to each other and opposite in phase. The net result is zero current. Since the current is zero, the circuit acts as a high resistance. The current supplied to a parallel-resonant circuit by the generator will be in phase with the generator voltage. The actual resistance of the parallel-resonant circuit can be obtained by measuring the voltage across it and dividing it by the current supplied by the generator. This effective resistance is known as the resonant resistance of the circuit.

There is a resonant voltage step-up in a series-resonant circuit. This is not the case, however, in a parallel-resonant circuit since the coil and the capacitor are connected in parallel. The current flowing between the coil and the capacitor is much higher than the current supplied by the generator. Therefore, we have a resonant current step-up in a parallel-resonant circuit. In a high-Q circuit of the latter type, the current flowing back and forth may be many times the line current.

VARYING R, L, C, AND F

Varying R.

You will remember that all coils have a certain amount of resistance which gives the effect of a resistor connected in series with the coil. This resistance can be changed by changing the size of the wire used to wind the coil while at the same time keeping the inductance of the coil constant.

If we use a circuit like the one shown in Fig. 15 to study the effect of varying the resistance in series with the coil, we will find that with the resistance set at a minimum, the coil current is equal to the capa-

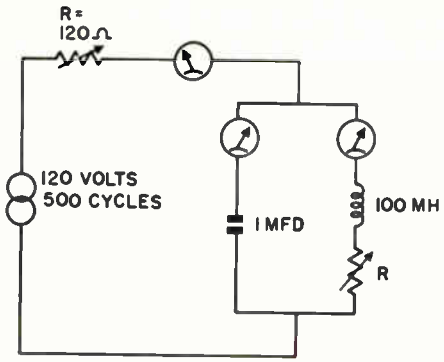


Fig. 15. Increasing R will increase the line current, which means that the resonant resistance of the resonant circuit has decreased.

citor current, and that the line current is very low. If we increase the value of the resistance R, the coil current will decrease slightly and the capacitor current will remain the same, but the line current will increase. This means that the resonant resistance of the parallel-resonant circuit must decrease in order for the line current to increase. From this we can see that the lower the coil resistance in a parallel-resonant circuit, the higher the resonant resistance and the lower the line current will be.

We mentioned earlier that once the current starts flowing back and forth in a parallel-resonant circuit, it will continue for a number of cycles even though the generator voltage may be removed. How quickly the current flowing in the circuit drops to zero depends upon the resistance in the circuit. If the resistance in the circuit is high, the energy in the circuit will be dissipated quickly in the resistor and the current will drop to zero in a few cycles. On the other hand, if the resistance in the circuit is very low, there will be very little energy lost each cycle and the back and forth action of the current may continue for a large number of cycles.

Varying C.

In the parallel-resonant circuit with a 100-millihenry coil, we will obtain resonance at 500 cycles when the capacity in parallel with the coil is 1 mfd. The line current will be minimal at this point.

If we set R to zero and try different values of capacitors in parallel with the coil, recording the line current for each capacitor, we could obtain the data to plot a curve like the one shown in Fig. 16. Notice that at resonance the line current drops to a low value. When the capacity is less than 1 mfd, the current rises until it is about .35 amp at zero capacity. Under these circumstances, with no capacity in parallel with the coil, the current flowing in the circuit will be limited by the inductive reactance of the coil and the amount of resistance in the circuit.

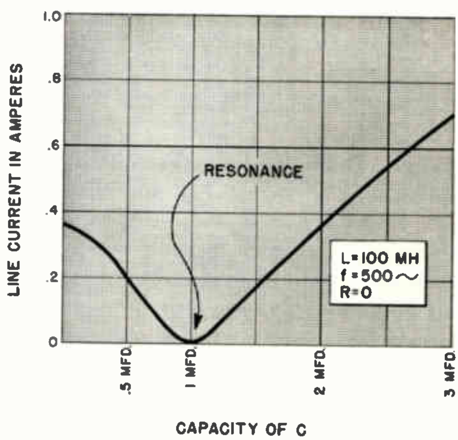


Fig. 16. How the line current varies when the capacity of the capacitor in a parallel-resonant circuit is varied.

When the capacity is reduced to zero, we have a circuit like the one in Fig. 2. However, as capacity is placed in parallel with the coil, the line current begins to drop until finally at resonance it is practically zero.

As the capacity is increased beyond 1 mfd, its reactance decreases, so you will find that the capacitor current will start to increase, the line current will increase, and the coil current will remain essentially unchanged. Here the increase in line current is due to the fact that the capacitor current becomes greater than the coil current and hence part of the capacitor current must be drawn from the line.

When the capacity is less than that required for resonance, the increase in line current is due to the fact that the coil current is greater than the capacitor current and part of the coil current must be drawn from the line.

It is interesting to note that when the capacity in the circuit is .5 mfd, the current is considerably higher than it was when the circuit was at resonance. This is because the capacitive reactance in the circuit is considerably higher than the inductive reactance. If the frequency applied to the circuit were increased, the inductive reactance would increase, and the capacitive reactance would decrease. By increasing the frequency by the correct amount, we could eventually reach the point where the inductive reactance would be equal to the capacitive reactance of the .5 mfd capacitor; once again we would have resonance.

Varying L.

If we put different coils in the circuit, while at the same time keeping the source frequency at 500 cycles,

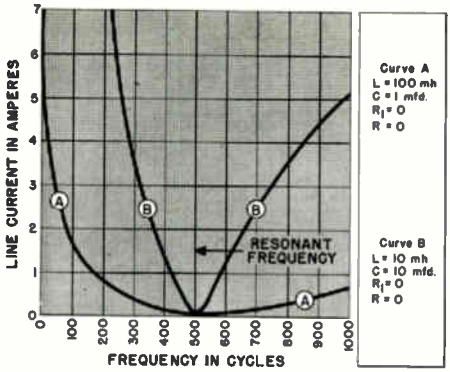


Fig. 17. How the line current varies when the frequency is varied in a parallel-resonant circuit.

we will find that we have an effect somewhat similar to the effect of changing the value of the capacitor. The line current will be increased when the inductance is made either too large or too small for resonance. If the inductance is below the value needed for resonance, the circuit will act exactly as it did when too low a capacity was used; when the inductance is too high for resonance, the circuit will act as it did when the capacity was too high.

Varying F.

In Fig. 17 we have a graph that shows how the line current will vary as the frequency applied to the resonant circuit is changed, if both resistors are set at zero. If we started with 0 cycles, which is dc, we would have a very high current. At this frequency there would be no current through the capacitor, and the inductive reactance of the coil would be zero. The only thing limiting the current flow in the circuit would be the resistance of the coil and of the leads used to connect it to the voltage source. However, as the frequency applied to the circuit increases, the line current drops until at 500 cycles the current is practically zero. As the frequency is increased beyond resonant frequency, the current will increase slowly. The increase in line current is due to the drop-off in the capacitive reactance of the capacitor. Current flowing through the coil will continue to decrease as the frequency is increased, because the inductive reactance of the coil will increase with the frequency.

At a frequency below the resonant frequency of the circuit, most of the current flows through the coil, and hence the parallel-resonant circuit acts as a coil. Right at resonance the circuit acts as a very high resistance, and above the resonant frequency the current flowing through the capacitor will be greater than the current flowing through the coil; hence, the resonant circuit will act as a capacitor.

The curve marked A in Fig. 17 represents an inductance of 100 mh and a capacity of 1 mfd. This circuit is resonant at 500 cycles, because the inductive reactance of the coil is equal to the capacitive reactance of the capacitor at this frequency. However, if we reduce the inductance to 10 mh and increase the capacity to 10 mfd, we will again have a situation where the inductive reactance is equal to the capacitive reactance at 500 cycles. In other words, a 10-mh coil will form a parallel-resonant circuit with a 10 mfd capacitor at a frequency of 500 cycles. The curve we would obtain by varying the frequency of the voltage applied to the parallel-resonant circuit made up of the 10-mh coil and the 10-mfd capacitor is represented by curve B in Fig. 17. Notice that the current rises much faster on both sides of resonance and drops to zero much more sharply than the curve for the 100 mh coil and the 1 mfd capacitor. We say that curve B is sharper than curve A. Curve A was obtained with one LC ratio: a 100 mh coil and a 1 mfd capacitor. Curve B was obtained with another LC ratio: a 10-mh coil and a 10-mfd capacitor. The LC ratio for curve A is higher than the LC ratio for curve B. A low LC ratio gives a sharp curve. This is an important thing to remember.

A low LC ratio is essential if a parallel-resonant circuit is to be used to separate signals having nearly the same frequency (for example, in radio receivers where stations operating on frequencies close together must be separated). If the resonant curve is sharp, we can tune in the desired signal and reject the undesired signals. However, if the resonant curve is broad, as the curve marked A in Fig. 17, it will be difficult to separate the undesired signals from the desired one.

The Q of a coil is another factor that will effect the sharpness of the resonance curves. A high-Q coil will yield a much sharper response curve than a low-Q coil.

In a series-resonant circuit, we obtain a sharp response curve with a high LC ratio. We have the opposite situation in a parallel-resonant circuit, however: we obtain a sharp curve with a low LC ratio.

SELF-TEST QUESTIONS

(j) How do you distinguish between a series-resonant and a parallel-resonant circuit?
(k) What does a parallel-resonant circuit act like at resonance?
(l) Does the generator supply a current of high value or of

low value to a parallel-resonant circuit?

(m) In circuits such as the one shown in Fig. 13, why will the voltage across a 120-ohm resistor be small?

(n) Does a current of high value or low value flow in the coil and capacitor in a parallel-resonant circuit?

(o) Will increasing the resistance of the coil in a parallel-resonant circuit cause the generator current to increase or to decrease?

(p) If a parallel-resonant circuit is used in a radio receiver to select one signal and reject others, do you want a high LC ratio or a low LC ratio?

Comparison of Series-Resonant And Parallel-Resonant Circuits

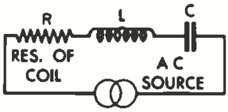
Series-resonant and parallel-resonant circuits are found in every radio and TV receiver and in many other pieces of electronic equipment. Resonant circuits are used in receiving equipment to separate the stations operating on different frequencies and in transmitting equipment in conjunction with vacuum tubes and/or transistors to generate radio frequency signals.

The chart shown in Fig. 18 compares and summarizes the important characteristics of series-resonant and parallel-resonant circuits. Notice that in many cases a series-resonant circuit is the exact opposite of a parallel-resonant circuit. Perhaps the most important characteristics of the two types are the resistance at resonance and the current at resonance. A series-resonant circuit acts as a low resistance at resonance and the current flowing through it will be at its maximum value. On the other hand the parallel-resonant circuit is exactly the opposite, so that at resonance it acts as a very high resistance and the line current flowing through it will be at its lowest value.

RESONANCE CURVES

In Fig. 17 you saw that we used a small inductance and a large capacity to obtain a sharp resonance curve with a parallel-resonant circuit. This gave us a low LC ratio. For a sharp resonant curve in a series-resonant circuit you should use a high LC ratio; in other words, you should use a large inductance and a small capacity. This is simply another example of the difference between series-resonant and parallel-resonant circuits. As a technician you will not be called upon to design a series-resonant or a parallel-resonant circuit, but the more you understand about the circuits the better you will be able to maintain the equipment for which you may be responsible.

SERIES-RESONANT CIRCUITS



1. The coil, the capacitor and the AC voltage source are all in series.

2. Resonance occurs when the reactance of L is equal to the reactance of C.

3. At resonance, source current is a *maximum* (very high).

4. At resonance, a series resonant circuit acts like *a resistor of low ohmic value.*

5. At resonance, the voltages across L and C are equal in magnitude but 180 degrees out of phase with each other.

6. At resonance, the same current flows through the entire circuit.

7. At resonance, the voltage across either L or C may be greater than that of the source, giving resonant voltage step-up.

8. At resonance, increasing the value of coil resistance R lowers the circuit current, thereby *lowering* the resonant voltage step-up.

9. Off resonance, the circuit acts like that part which has the *higher* reactance.

   a. Increasing C above its at-resonance value makes the circuit act like a coil.

   b. Reducing C below its at-resonance value makes the circuit act like a capacitor.

   c. Increasing L above its at-resonance value makes the circuit act like a coil.

   d. Reducing L below its at-resonance value makes the circuit act like a capacitor.

   e. Applying a *higher* frequency than the resonant one makes the circuit act like a coil.

   f. Applying a *lower* frequency than the resonant one makes the circuit act like a capacitor.

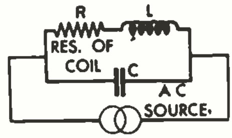
10. The product LC is constant for any given resonant frequency.

11. Increasing L or increasing C lowers the resonant frequency.

12. Decreasing L or decreasing C raises the resonant frequency.

13. The Q factor of the circuit is essentially equal to the coil reactance divided by the AC resistance of the coil.

PARALLEL-RESONANT CIRCUITS



1. The coil, the capacitor and the AC voltage source are all in parallel.

2. Resonance occurs when the reactance of L is equal to the reactance of C.

3. At resonance, source current is a *minimum* (very low).

4. At resonance, a parallel resonant circuit acts like *a resistor of high ohmic value.*

5. At resonance, the voltages across L, C and the source are all the same in magnitude and phase.

6. At resonance, the currents through L and C are essentially equal in magnitude but are 180 degrees out of phase.

7. At resonance, the current through either L or C is greater than the source current, giving resonant current step-up.

8. At resonance, increasing the value of coil resistance R increases line current, thereby *lowering* the resonant current step-up.

9. Off resonance, the circuit acts like that part which has the *lower* reactance.

   a. Increasing C above its at-resonance value makes the circuit act like a capacitor.

   b. Reducing C below its at-resonance value makes the circuit act like a coil.

   c. Increasing L above its at-resonance value makes the circuit act like a capacitor.

   d. Reducing L below its at-resonance value makes the circuit act like a coil.

   e. Applying a *higher* frequency than the resonant one makes the circuit act like a capacitor.

   f. Applying a *lower* frequency than the resonant one makes the circuit act like a coil.

10. The product LC is constant in any given resonant frequency.

11. Increasing L or increasing C lowers the resonant frequency.

12. Decreasing L or decreasing C raises the resonant frequency.

13. The Q factor of the circuit is essentially equal to the coil reactance divided by the AC resistance of the coil.

Fig. 18. Comparison of series-resonant and parallel-resonant circuits.

DISTINGUISHING BETWEEN SERIES AND PARALLEL-RESONANT CIRCUITS

Sometimes, it is not easy to distinguish between a series-resonant and a parallel-resonant circuit. In Fig. 19A we have shown a series-resonant circuit. A series-resonant circuit is a resonant circuit in which the source voltage has been applied to the coil and capacitor in series. There is no doubt that this is a series-resonant circuit.

In Fig. 19B we have shown a parallel-resonant circuit. The parallel-resonant circuit is a resonant circuit in which the source voltage has been applied to the coil and capacitor in parallel. Again, it is easy to see that this is a parallel-resonant circuit.

Fig. 19C shows two resonant circuits that again look like two parallel-resonant circuits. Here, the secondary is inductively coupled to the primary. Let's look at the primary first. The voltage source is applied to the coil and capacitor in parallel; there is no doubt that the primary is in a parallel-resonant circuit. But how about the secondary? Since the coil and capacitor are connected in parallel you might jump to the conclusion that this is a parallel-resonant circuit, too. Actually, this has no bearing--how the voltage is applied to the circuit determines whether the circuit is a series-resonant or parallel-resonant circuit.

The voltage is induced in the secondary. Actually, some voltages are being induced in each turn of the coil and they act as if they are connected in series, so that the total voltage induced in the secondary is the sum of the voltages induced in each turn. We can compare this to a number of small generators connected in series with the various turns of the coil, and the coil might look like Fig. 19E.

Thus, the voltage induced in the coil is actually applied in series with the turns of the coil rather than in parallel with the coil and the capacitor and could be represented by Fig. 19D (which is the same as Fig. 19A). Therefore, the secondary of the transformer shown in Fig. 19C is a series-resonant circuit and not a parallel-resonant circuit.

You will run into this type of double-tuned circuit in many pieces of electronic equipment. It is often used between two stages in a radio receiver or a television receiver as shown in Fig. 20. Here the primary is connected between the plate of one tube and B+. The secondary is connected between the grid and the cath-

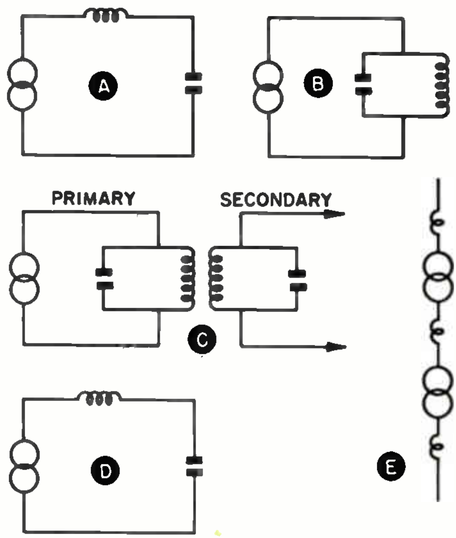


Fig. 19. A series-resonant circuit is shown at A, and a parallel-resonant circuit is shown at B. In C, the primary of the transformer and the capacitor across it form a parallel-resonant circuit, and the secondary of the transformer and its capacitor form a series-resonant circuit.

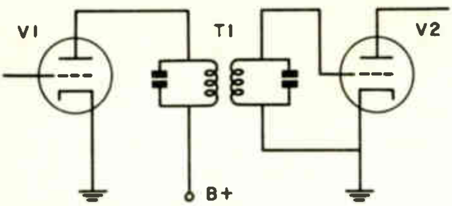


Fig. 20. The primary of T1 is a parallel-resonant circuit, but the secondary is a series-resonant circuit.

ode of the second tube. The tube marked V1 really acts as a generator and supplies the ac signal across the primary of the transformer. A signal is induced in series with the secondary winding, because the secondary is inductively coupled to the primary. Because the coil and capacitor in the secondary form a series-resonant circuit, there is a high current flow with resulting resonant voltage step-up so that this stepped up voltage is applied between the grid and the cathode of V2.

The resonant voltage step-up that occurs in the secondary winding of T1 is quite important. Actually, the primary and secondary windings of T1 will have the same number of turns in most cases. Therefore, you would expect the voltage induced in the secondary to be approximately equal to the voltage across the primary. This is what would happen if we simply had a transformer in the circuit. However, since there are capacitors across each coil and since each circuit is a resonant circuit we have the resonant voltage step-up which occurs in the secondary winding due to the high current flowing in the series-resonant circuit. Thus, there is actually a step-up in voltage occurring in the transformer even though the turns-ratio may be one to one. This means that the signal voltage available between the grid and cathode of V2 will be considerably greater than the voltage between the plate and ground of V1.

SELF-TEST QUESTIONS

(q) Explain the difference between the current flowing in the coil and capacitor in a series-resonant circuit and the current flowing in the coil and capacitor in a parallel-resonant circuit.

(r) Explain the difference between the voltage across the coil and capacitor in a series-resonant circuit and the voltage across the coil and capacitor in a parallel-resonant circuit.

(s) What is the difference between the generator current in a series-resonant circuit and in a parallel-resonant circuit?

(t) How can the voltage across the coil or capacitor in a series-resonant circuit be greater than the source voltage?

(u) What happens if we increase the value of inductance or capacitance in a resonant circuit?

(v) Will the voltage across the coil or the capacitor in a high-Q series-resonant circuit be greater than the voltage across the coil or capacitor in a low-Q series-resonant circuit?

(w) Reducing L below its at-resonance value in a series-resonant circuit makes the circuit act as a capacitor; reducing L below its at-resonance value in a parallel-resonant circuit makes the circuit act as a coil. Explain why this happens.

How Resonant Circuits Are Used

Resonant circuits have many applications. In this section of the lesson we will look into some of the more common uses of resonant circuits in radio and TV. These uses are important because they demonstrate how resonant circuits are used to select one signal from a number of signals of different frequencies.

SELECTING A DESIRED SIGNAL

The antenna connected to a radio or a TV receiver picks up signals from a large number of radio and TV stations. Even the antennas installed especially for television receivers will pick up a certain amount of signal from radio broadcast-band stations. This happens even though the antenna is designed for operation on a much higher frequency than that of the broadcast-band station. Therefore, some means must be provided inside the receiver to select the desired signal and reject the unimportant one. Resonant circuits are used for this purpose.

In Fig. 21 we have shown the input circuit of a radio receiver. The

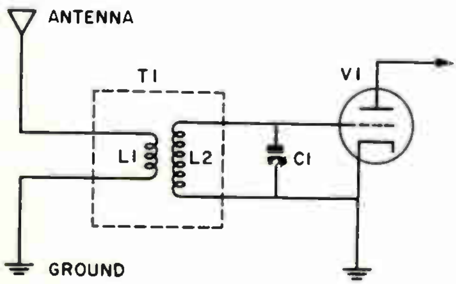


Fig. 21. The input circuit of a typical radio receiver.

signals picked up by the antenna cause a current to flow through the primary winding L1 of T1. T1 is called an antenna coil or transformer because the signals from the antenna are applied to this coil. The secondary L2 of T1 is inductively coupled to the primary so that the current flowing in L1 sets up a magnetic field which cuts L2 and induces a voltage in it. Remember that when a voltage is induced in a coil in this way, there is a certain amount of voltage induced in each turn of the coil. The voltage induced in the coil acts like a number of generators connected in series with the coil. Thus L2 and capacitor C1 form a series-resonant circuit at some frequency within the broadcast band.

Notice the symbol used for the capacitor C1. This symbol indicates that the capacitor is variable. Thus, by changing the setting of C1, the frequency at which the combination of L2 and C1 is resonant can be changed.

Let us suppose that the antenna is picking up two signals of equal amplitude or strength, one having a frequency of 500 kc and the other having a frequency of 800 kc. If the combination of C1 and L2 is resonant at 800 kc (C1 and L2 thus forming a series-resonant circuit), there will be a high 800-kc current through L2 and C1, with resultant step-up voltages appearing across L2 and C1. These voltages are applied between the grid and the cathode of V1 to be amplified by this tube.

At the same time there is a 500-kc signal being picked up by the antenna. This will flow through the primary of transformer T1 and will induce a certain voltage in L2. Since the combination of L2-C1 is not resonant at 500 kc, the impedance of this series circuit will be much higher at 500 kc than it was at 800 kc. This means that the 500-kc current flowing through the series circuit will be low so that voltage developed across L2 and across C1 by this current will be low. Therefore, the 500-kc signal applied between the grid and the cathode of V1 will be much lower in amplitude than the 800-kc signal. Thus, although one resonant circuit is not able to reject the 500-kc signal completely, the amplitude of this signal (when applied between the grid and the cathode of V1) is lower than the amplitude of the desired 800-kc signal that is applied to this tube.

Better selectivity can be obtained in a receiver by using several resonant circuits. If each circuit is tuned to the desired frequency, the difference in signal strength between the desired and the undesired signals will become greater. If enough resonant circuits are used the only signal actually heard in the output of the receiver will be the signal from the desired station.

I-F Transformers.

Most modern radio and television receivers use the superheterodyne principle. In the superheterodyne receiver the signal is picked up by the antenna and fed to a stage called a mixer or first detector. Here, the signal is mixed with a signal generated by the oscillator stage. The two signals mixed together produce two new signals, one equal to the sum of the two frequencies and the other equal to the difference of the two. Both new signals contain the modulation on the original signal. In a superheterodyne receiver we use the difference-frequency signal. In the output circuit of the mixer stage we use a transformer called an intermediate frequency transformer (usually abbreviated as an i-f transformer). This transformer is tuned to resonance at the difference frequency. One or more amplifier stages (called i-f amplifiers) are used to amplify the difference signal. I-F transformers are used between the mixer and the various stages in the i-f amplifier and between the last i-f amplifier stage and the stage called the second detector. The latter separates the audio or picture signals from the rf carrier. Resonant circuits are used in i-f transformers.

The schematic of a circuit used between the mixer and first i-f stage in a superheterodyne receiver is shown in Fig. 22. The tube marked V1 is the mixer; the tube marked V2, the i-f tube.

You have already seen this type of circuit earlier in the lesson. You know that the primary of T1 is a parallel-resonant circuit because the tube acts as a generator and applies the signal in parallel with the resonant circuit. This parallel-resonant circuit acts as a high resistance at the resonant frequency and the voltage developed by the tube will be high. At frequencies

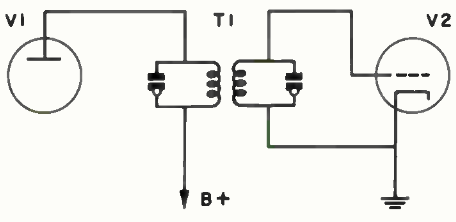


Fig. 22. Coupling between the mixer and the i-f tube in a superheterodyne receiver.

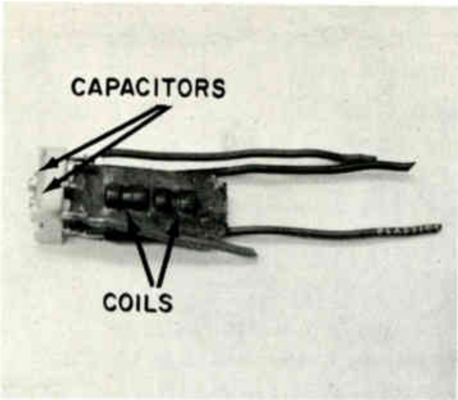


Fig. 23. The inside of an i-f transformer.

other than the resonant frequency, the primary circuit of T1 does not act as a high resistance; as a matter of fact, it acts as a fairly low impedance, so the voltage developed at these frequencies by V1 is low.

The primary of T1 is inductively coupled to the secondary winding so that the secondary circuit is a series-resonant circuit. Again at the resonant frequency, a high current flows and there is considerable resonant voltage step-up across the coil and across the capacitor. These stepped-up voltages are applied between the grid and the cathode of the i-f stage.

A typical i-f transformer is shown in Fig. 23. Notice that the two coils are placed near each other so that the primary and secondary are inductively coupled together. At the top of the transformer are two trimmer capacitors. The adjusting screws on these capacitors can be reached through holes in the top of the i-f transformer can or shield and can be adjusted for exact resonance after they have been installed in the circuit.

A modern superheterodyne receiver uses at least two i-f transformers like the one shown in Fig. 23. The selectivity of two transformers in conjunction with the selectivity obtained in other circuits will make the receiver selective enough so that it will pick up the desired signal even in the crowded broadcast band and in most cases reject the signals from undesired stations.

High-Frequency Circuits.

In some high frequency applications you might find a circuit like the one shown in Fig. 24. The symbol beside the coil indicates that the coil has a slug which can be adjusted in and out of the coil. This will change the inductance of the coil.

Although no capacitor is shown in the circuit, the circuit is actually a parallel-resonant circuit. The tube has a certain capacity between plate and ground, and this capacity will be in parallel with the coil. At high frequencies this capacity, along with the coil, is all that is needed to form a resonant circuit. Circuits of this type are frequently found in television receivers.

In some applications the coil may consist of less than one turn of a flat ribbon-type material such as

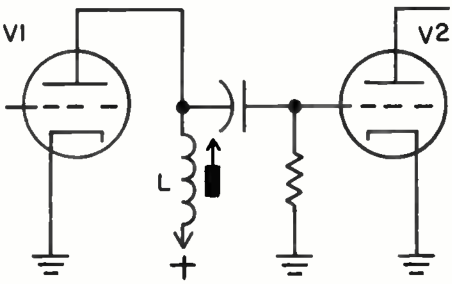


Fig. 24. Coil L in the plate circuit of V1 along with the circuit capacities form a parallel-resonant circuit.

shown in Fig. 25. This type of coil is used in UHF TV circuits and although the inductance in the circuit is extremely small, due to the fact that the circuit must operate at several hundred megacycles, this inductance is all that is required. As a matter of fact, in resonant circuits designed for UHF operation, the problem is not in getting the needed inductance and capacitance, but rather in keeping the inductance and capacitance low enough to produce resonance at the ultra-high frequency desired.

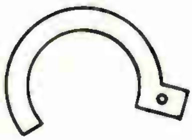


Fig. 25. A single turn of flat ribbon is all the coil that is needed in UHF circuits.

HOW DIFFERENT TYPES OF FILTERS ARE USED

Another important use of resonant circuits is in the design of filters. There are three different types of filters that you are likely to encounter as a technician. The explanation of exactly how each type works is rather complex, and since you need not know how each type of filter operates, we will not go into an explanation here. However, it is important that you know how the different types of filters are used.

Low-Pass Filters.

In Fig. 26 we have shown a schematic diagram of a low-pass filter, which is a filter that will allow signals below a certain frequency to pass through it with little or no attenuation, which means weakening of signals above this specific frequency. For example, if a low-pass filter is designed to pass frequencies below 10 megacycles, it will pass all frequencies from zero cycles per second, which is dc, up to 10 mc with little or no opposition. However, a signal with a frequency of 15 mc, or 25 mc, or in fact any frequency above 10 mc, will encounter a great deal of opposition in going through the filter.

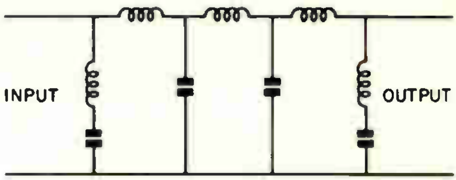


Fig. 26. A low-pass filter.

High-Pass Filters.

A high-pass filter is designed to cut off all frequencies below a certain frequency and allow signals above this frequency to pass through with little or no attenuation. A schematic of a typical high-pass filter is shown in Fig. 27.

High-pass filters are often used on TV receivers to eliminate interference from stations operating on frequencies below the television channel. High-pass filters designed for this purpose are available commercially.

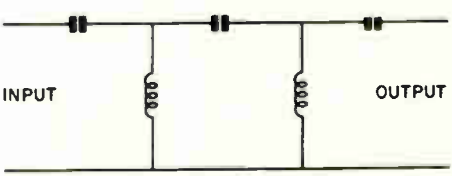


Fig. 27. A high-pass filter.

Band-Pass Filters.

Another type of filter is shown in Fig. 28. This type of filter, called a band pass filter, allows a certain band of frequencies to pass through it with little or no attenuation but offers considerable opposition to signals above and below the frequency of the band to be passed.

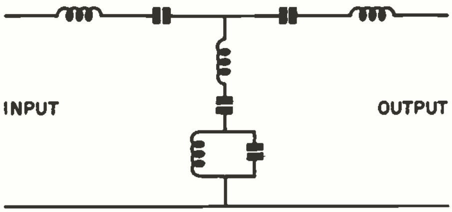


Fig. 28. A band pass filter.

For example, if a band pass filter is designed to have a band width of 2 megacycles, and the center of the pass band is 10 megacycles, then the band pass filter will pass frequencies from 9 mc to 11 mc with little or no attenuation. However, a signal having a frequency of 7 mc, which is below the pass band, or a signal having a frequency of 15 mc, which is above the pass band, will encounter considerable opposition going through the band pass filter.

SUMMARY

In this section of the lesson we have covered a few of the most important uses of resonant circuits. Resonant circuits are used in radio and TV receivers to select one desired signal and reject others. Both series-resonant and parallel-resonant circuits are used in the input stages of a receiver. They are also used between the mixer and i-f stages and between the i-f stage and the second detector.

Resonant circuits are used in filters. A low-pass filter is a filter which will pass frequencies below a certain frequency with little or no attenuation but offers high attenuation to signals above this frequency. A high-pass filter is a filter that will offer little or no attenuation to signals above a certain frequency, but offers high opposition or attenuation to signals below this frequency. A band pass filter will pass a certain band of frequencies, but attenuate signals either above or below the band of frequencies which it is designed to pass.

SELF-TEST QUESTIONS

(x) Is the resonant circuit made up of C1 and L2 in Fig. 21 a series-resonant circuit or a parallel-resonant circuit?
(y) What is a low-pass filter?
(z) What is a high-pass filter?
(aa) What is a band pass filter?
(ab) What type of circuit is used in making up filters?

RC Circuits

Another type of circuit that is extremely important in electronics work is the RC circuit, so called because it contains resistance and capacity. There are several types of RC circuits.

An RC circuit is used as a coupling circuit. This type of circuit is designed to pass a signal through it without changing the shape of the signal. Circuits of this type are used where an ac signal is mixed with dc. RC coupling circuits are widely used in the audio sections of radio and television receivers between the various stages. An RC coupling circuit used between two tubes is shown in Fig. 29A and an RC coupling circuit used between two transistor stages is shown in Fig. 29B. The purpose of the coupling circuit in each case is to pass the signal from one stage to the other without changing the shape of it; at the same time the circuit keeps the operating voltages from one stage out of the following stage.

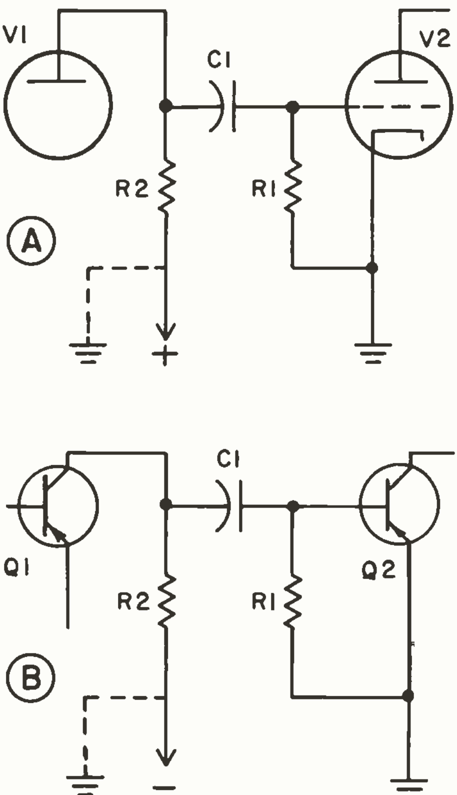


Fig. 29. RC coupling circuits used between tube and transistor stages.

Another type of RC circuit is designed specifically to change the shape of the signal applied to it. This type of circuit is used because the signal being fed through the RC circuit may be used to control the following stage. The shape of the signal may not be the best possible shape to control the stage; by means of a suitable RC circuit, the shape of the wave can be altered. In this section we will study both coupling and wave-shaping circuits, and you will see what they look like, how they work, and where each type is found.

RC COUPLING CIRCUITS

You will run into RC circuits most frequently between stages where they are used to feed the signal from one stage to the next. Typical RC coupling circuits are shown in Fig. 29. The RC coupling circuit in both cases consists of C1 and R1. These are the two components that you will be most concerned with in determining the characteristics of this type of circuit.

The tube V1 in Fig. 29A and the transistor Q1 in Fig. 29B act essentially like an ac generator in RC coupling circuits. One end of resistor R2 is connected to the tube in A and to the transistor in B. The other end is connected to the power supply. However, as far as the ac signal is concerned the end of R2 which is connected to the power supply is in effect connected to ground, because there is a large capacity in the power supply output connected between B+ and ground. This capacitor is so large that it has a very low reactance at all signal frequencies. Therefore the first stage, which is V1 in Fig. 29A and Q1 in Fig. 29B, supplies the voltage to resistor R2. The resistor acts like it is connected between the tube or transistor and ground.

An equivalent circuit of the coupling network is shown in Fig. 30. Here we have represented the tube or transistor as a generator with R2 connected across it. Notice that C1 and R1 are connected in series with each other and this combination is connected in parallel with the resistor R2. The purpose of the coupling network C1-R1 is to feed the signal that is across R2 to the following stage, V2 in Fig. 29A and Q2 in Fig. 29B. If the reactance of C1 is low enough, it will act as a short circuit at signal frequencies so R1 will in effect be connected in parallel with R2. When this situation exists, all the voltage available at the generator output will appear across R1. In other words the voltage appears between the grid and cathode of V2 in Fig. 29A and between the base and emitter of the transistor in Fig. 29B.

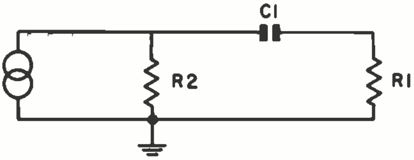


Fig. 30. The equivalent circuit of the RC coupling network shown in Fig. 29.

Capacitor C1 and resistor R1 are in series and they form a voltage divider network. Part of the voltage developed across R2 will be dropped across C1 and part of it across R1. The more voltage there is across R1 the more voltage we have available to drive the second stage. It is therefore important that the reactance of C1 be kept as low as possible in comparison with the resistance of R1. However, regardless of how large the capacitor C1 is, its reactance will eventually become high enough at some low frequency so that an appreciable part of the voltage developed across R1 is lost across C1. At some frequency the reactance of C1 will be equal to the resistance of R1. When this situation occurs 70.7% of the voltage appearing across R2 will be present across R1. You might expect only 50% of the voltage to appear across R1. However, 70.7% is correct because the voltage across R1 is not in phase with the voltage across C1. If the frequency is made still lower, then the percentage of voltage appearing across R1 will be lower.

When the voltage across R1 drops to 70.7% of the generator output, the current flowing through R1 will be only 70.7% of the maximum it would be with the full voltage across R1. When the voltage and current across R1 drop to 70.7%, the power across R1 will decrease to 50%. This point is called the "half-power" point. The amplifier is considered satisfactory in most cases as long as the output does not drop below this point.

Since C1 and R1 form a voltage divider network, we can keep the frequency at which the reactance of

C1 becomes a problem to a fairly low value by making R1 as large as possible. In the tube circuit shown in Fig. 29A, R1 is called a grid leak. Its value is relatively unimportant and a comparatively large resistor can be used. Therefore, even with a fairly small capacitor for C1, the resistance of R1 does not have much importance except at very low frequencies. On the other hand, making R1 too large will upset the operating voltages in the transistor in the circuit shown in Fig. 29B. In fact, the transistor itself affects the circuit so that there is a maximum value resistor that can be used; it is much lower than that used in the tube stage. The capacitor C1 must be of a much higher capacity in the transistor circuit than in the tube circuit to keep its reactance from becoming high enough to drop an appreciable percentage of the voltage.

Capacitor values from .01 mfd to .05 mfd are typical in tube circuits such as in Fig. 29A. In transistor circuits the value of C1 will often be 10 mfd or more.

The importance of avoiding this voltage division can be seen if you consider what happens to signals of different frequencies when they are amplified by the amplifier. If a signal voltage of 1 volt and a frequency of 1000 cycles appears across R2, almost the full 1 volt will appear across R1. But, at a frequency of 100 cycles, somewhat less than 1 volt will appear across R1; at a frequency of 10 cycles, even less will appear across R1. This means that the amplifier will not amplify signals of different frequencies equally. When this situation exists, we say we have frequency distortion. In the average radio receiver, a small amount of frequency distortion is not objectionable, but in high-fidelity equipment and in TV equipment, this type of distortion must be kept at a minimum if satisfactory results are to be obtained.

For the present the important thing you should remember is that an RC coupling circuit has a long time constant. Remember that capacitor C1 and resistor R1 are in series, but that at most frequencies the reactance of C1 is so low compared to the resistance of R1 that it acts as a short circuit and can be ignored. However, at very low frequencies the reactance of this capacitor is appreciable. We will discuss this situation in more detail when you study vacuum tubes and amplifiers.

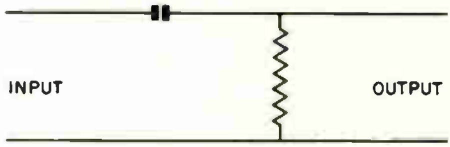


Fig. 31. An RC differentiating circuit.

RC DIFFERENTIATING CIRCUITS

Differentiating circuits are found in TV receivers and in many other pieces of electronic equipment. Fig. 31 shows a circuit of this type -- notice that it looks like a coupling circuit. However, the latter has a long time constant, whereas a differentiating circuit has a short time constant. Because differentiating circuits are normally used with pulses, you should know that a pulse is a variation of a quantity whose value is normally constant.

Fig. 32A illustrates a typical sine wave. Remember that the voltage starts at zero, builds up to a maximum and drops back to zero.

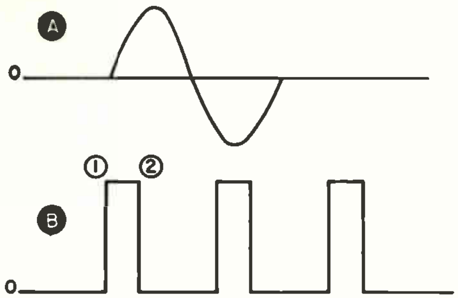


Fig. 32. A single cycle of a sine wave is shown at A, and a group of three pulses is shown at B.

A series of three positive-going pulses is shown in Fig. 32B. Notice that the signal voltage is zero, jumps instantly to its maximum value at point 1, remains constant from point 1 to point 2, and finally drops to zero again at point 2. Let's see what will happen if a pulse of this type is fed to a differentiating circuit such as the one in Fig. 31.

The value of R and of C in the differentiating circuit are chosen with a short time constant so that the capacitor charges and discharges quickly. As the leading edge of the pulse (which we have marked as 1) hits the capacitor there is an immediate current flow through the resistor to charge the capacitor. Since the capacitor has no charge on it at this instant, the current flow as well as the voltage developed across the resistor will be high. As the pulse maintains its constant value, the capacitor charges rapidly, the current flowing in the capacitor decreases and the voltage developed across the resistor falls off. Finally, the capacitor is fully charged, no additional current is flowing and the voltage across the resistor drops to zero.

The capacitor must discharge when the pulse drops from its peak amplitude at point 2 to zero. At the instant the applied voltage pulse disappears the capacitor is charged, although there is no longer a pulse across it. Then it starts to discharge through the resistor and a high current flows through the resistor. Instantly, there is a high voltage across it. The current flows in the opposite direction from that in which it flowed while the capacitor was charging, so the pulse appears in the opposite direction. As the capacitor becomes discharged, the current flowing drops gradually to zero so that the voltage appearing across the resistor also drops to zero.

The reaction of a differentiating circuit to a series of pulses is shown in Fig. 33. At A we have shown the pulse and below it at B the output voltage that will be obtained across the resistor as the pulse varies to a maximum and drops back to zero again.

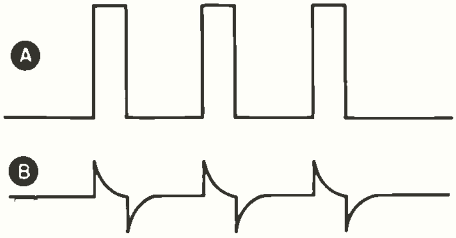


Fig. 33. Input and output signals applied to a differentiating circuit.

Differentiating circuits work in this way because they have a short time constant. The capacitor is able to charge and discharge rapidly, and therefore a double-pointed pulse as shown in Fig. 33 is obtained when a pulse is applied to the input of this type of circuit.

RC INTEGRATING CIRCUITS

A typical RC integrating circuit is shown in Fig. 34. Notice that the parts are connected in a way opposite from that in which they were connected in the differentiating circuit, and that a long time constant rather than a short time constant is used.

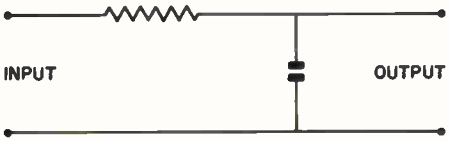


Fig. 34. An integrating circuit.

Fig. 35 illustrates the action of an integrating circuit. The capacitor begins to charge as the first pulse strikes the circuit; it starts to discharge after the first pulse passes but does not discharge completely before the second pulse arrives due to the long time constant. After the second pulse the capacitor starts to discharge again until the third pulse arrives -- then it charges still further. An integrating circuit is thus able to sum up or add a series of pulses to give one pulse in the output. Integrating as well as differentiating circuits are used in TV receivers and in many other pieces of electronic equipment.

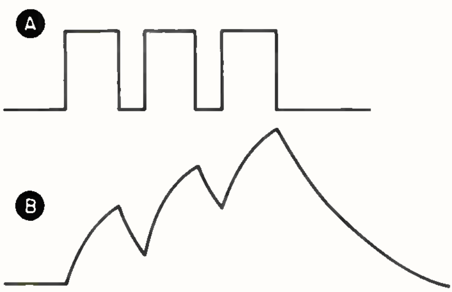


Fig. 35. Integrating circuit action on a series of pulses.

SUMMARY

In this section of the lesson you found that RC circuits can be used in several ways. They can be used as coupling circuits to feed a signal from the plate of one tube to the grid of the next tube. When they are used in this way the purpose is to pass the signal from one tube to the other without distorting or changing it in any way.

RC circuits are also used as differentiating circuits. A differentiating circuit is a circuit that develops a sharp positive and negative pulse from a single pulse. If the pulse supplied to the circuit is a positive-going pulse, which is what we call the pulse shown in Fig. 32B, then the output will be a positive pulse followed by a negative pulse. If the input signal is a negative-going pulse, then the output will be a sharp negative pulse followed by a sharp positive pulse.

An integrating circuit is a circuit that has a long time constant and adds together a number of separate pulses to produce one large pulse in the output.

SELF-TEST QUESTIONS

(ac) What is the purpose of an RC coupling circuit?

(ad) Is the reactance of the coupling capacitor likely to become a problem at high, low, or medium frequencies?

(ae) What do we mean by the half-power point?

(af) What do we mean by frequency distortion?

(ag) What is a differentiating circuit?

(ah) What is an integrating circuit?

LOOKING AHEAD

In this lesson you have seen how resistors, coils, and capacitors are used together to form a number of different types of circuits. By this time you have probably realized the importance of these three components. Before leaving this lesson it is worthwhile to stop and consider the fact that we have all three of these quantities in every circuit.

Even a piece of straight wire has a certain amount of resistance, a certain amount of capacity between it and nearby objects, and also a small amount of inductance. If the frequency of the signal running through the wire is high enough, even these small amounts of resistance, capacity, and inductance may be large enough to merit consideration.

We have brought up this point now because you will soon be working on, repairing, and replacing circuits in equipment which operates at very high frequencies. Because the resistance, capacity, and inductance in circuits operating at these high frequencies is so important, replacement parts should be put as closely as possible in the position occupied by the original part.

In the following lesson you will study additional components that will be important in your electronics career: vacuum tubes, transistors, complete stages and signals. You will also learn how the value of parts used in a circuit affects the performance of the stage.

There are a number of schematic diagrams in this lesson. If you study carefully the diagrams which appear in earlier lessons, complex diagrams in later lessons will not present a problem for you. Difficulties will occur if you wait until later lessons to trace circuits on complex diagrams.

ANSWERS TO SELF-TEST QUESTIONS

(a) A resonant circuit is one in which the inductive reactance cancels the capacitive reactance.

(b) We mean that the rms value or the effective value of the ac current flowing in the circuit is 1 amp. The ac current would have the same heating effect as 1 amp of dc. Remember that an ac current actually drops to 0 twice each cycle and reaches peak values approximately 1.4 times the effective or rms value.

(c) We cannot add the voltages because they are not in phase. The voltage across the resistor will be in phase with the current, whereas the voltage across the coil will lead the current by 90°. We must add these two voltages by means of vectors.

(d) In the expression "j50", the j indicates a reactive component; the 50 (for ohms) represents a reactance rather than a resistance. The plus sign in front of the j means that the reactance is inductive. A minus sign in front of the j indicates capacitive reactance.

(e) The voltage across the parts in an RC circuit will depend upon the parts themselves. If the reactance of the capacitor is greater than the resistance of the resistor then the voltage across the capacitor will be greater than the voltage across the resistor. On the other hand, if the resistance of the resistor is higher than the reactance of the capacitor, the voltage across the resistor will be greater than the voltage across the capacitor.

(f) Across the resistor the voltage will be in phase with the current; across the coil it will lead the current by 90°; across the capacitor it will lag the current by 90°.

(g) The Q of a coil is equal to the inductive reactance of the coil divided by the resistance of the coil. It is indicative of the worth of the coil, which is supposed to have inductive reactance with little or no resistance. Since the coil is wound with wire, however, it will have some resistance. The coil will act more like a resistor as this resistance increases. Therefore, a high-Q coil is better than a low-Q coil.

(h) The current will reach its maximum value when the inductive reactance of the coil cancels the capacitive reactance of the capacitor. When this happens we have a series-resonant circuit.

(i) Reducing the resistance in a series-resonant circuit will cause the current flowing in the circuit to increase. A higher resonant voltage will appear in turn across the coil and the capacitor. The voltage which appears across the resistor will remain the same because the increase in current will be counteracted by the reduction in resistance. Thus, the entire generator voltage will appear across the resistor.

(j) The distinction between a series-resonant circuit and a parallel-resonant circuit lies in the way in which the voltage is applied to the coil and the capacitor. If it is applied to the coil and the capacitor in series, the circuit is series-resonant; if it is applied to the coil and the capacitor in parallel, the circuit is parallel-resonant.

(k) A high value resistance.

(l) The generator connected across a parallel-resonant circuit will supply a current of low value because the high resistance of this type of circuit limits the current which can flow.

(m) The voltage across the resistor will be small because the resistor is in series with the parallel-resonant circuit. Most of the voltage will be dropped across the higher resistor. A parallel-resonant circuit has a very high resistance at resonance and most of the voltage will be dropped across it. Consequently, there will be very little voltage across the 120-ohm resistor.

(n) In a parallel-resonant circuit, a high value current flows in the coil and in the capacitor. These two currents are 180° out of phase. Energy stored in the capacitor flows out of the capacitor into the coil, where it is stored, and then flows from the coil back into the capacitor. This back-and-forth flow of current reaches a very high value. Even though the current actually supplied by the generator is low, it is enough to make up for losses in the resonant circuit due to resistance in the coil and in the wires connecting the coil and the capacitor together.

(o) Increasing the resistance of the coil in a parallel-resonant circuit causes the generator current to increase. More losses occur as a result of increased resistance, and the generator supplies more current to make up for these losses.

(p) A low LC ratio gives a sharper curve such as curve B in Fig. 17. This type of curve is required in order to select one station and reject another. A broad curve such as curve A of Fig. 17 would be unsuitable because stations operating close to the desired station would not be rejected.

(q) In a series-resonant circuit the same current flows in the coil and in the capacitor. As a matter of fact, since the generator, resistance, coil and capacitor are all in series in a series-resonant circuit, the same current must flow through all these components. On the other hand, in a parallel-resonant circuit the current through the coil and capacitor are essentially equal in magnitude, but they are 180° out of phase.

(r) In a series-resonant circuit, the voltage across the coil will be equal to but 180° out of phase with the voltage across the capacitor. In a parallel-resonant circuit the coil and capacitor are connected in parallel and the voltage across the two will therefore be the same.

(s) The generator current in a series-resonant circuit will be very high because this circuit acts as a low resistance. The generator current in a parallel-resonant circuit will be very low because this circuit acts as a high resistance.

(t) The inductive reactance of the coil in a series-resonant circuit cancels the capacitive reactance of the capacitor. Therefore the only factor that limits current flow in the circuit is the resistance in the circuit. This results in a very high current flow. The current flowing through the coil and through the capacitor produces a voltage drop across these components which will be equal to the product of the current times the reactance of the particular part. This product may be greater than the source voltage. The voltage across the coil and across the capacitor are 180° out of phase so that they cancel each other and the entire generator voltage will appear across the resistance in the circuit.

(u) Increasing the value of L or C in a resonant circuit will reduce the resonant frequency.

(v) Yes, the voltage across the coil and capacitor in a high-Q series-resonant circuit will be greater than the voltage across the coil and capacitor in a low-Q series-resonant circuit. This is due to the fact that a higher current will flow in a high-Q circuit which, all other factors being equal, will produce a greater voltage across the coil and across the capacitor.

(w) When we reduce L below its at-resonance value in a series-resonant circuit, the inductive reactance will be less than the capacitive reactance. Therefore the net reactance in the circuit will be capacitive - in other words the inductive reactance cannot completely cancel out the capacitive reactance. It will subtract from it but there will still be capacitive reactance left over, and the circuit will act as a capacitor. The current flowing through the circuit will lead the voltage. On the other hand, when we reduce L below its at-resonance value in a parallel-resonant circuit, the inductive reactance will be less than the capacitive reactance and more current will flow through the lower reactive branch. The net result will be that the capacitive current cannot completely cancel out the inductive current. Therefore, the circuit will act as inductance and the voltage will lead the current.

(x) The resonant circuit is a series-resonant circuit because the voltage is induced in series with the turns of L2. Therefore, the voltage is applied in series with the coil and the capacitor.

(y) A low-pass filter is a filter designed to pass signals below a certain frequency and reject all signals above that frequency.

(z) A high-pass filter is a filter designed to pass all signals above a certain frequency and reject signals below that frequency.

(aa) A bandpass filter is a filter designed to pass a certain band of frequencies with little or no attenuation. It will reject or offer considerable opposition to frequencies above and below the band it is designed to pass.

(ab) Various combinations of series-resonant and/or parallel-resonant circuits are used in making up filters.

(ac) An RC coupling circuit is used to transfer a signal from one stage to another without changing the shape of the signal.

(ad) At low frequencies.

(ae) The half-power point is the frequency at which 70.7% of the voltage appears across the resistor in an RC coupling circuit. At this frequency the current will also have dropped to 70.7% of its maximum value so that the power will be down 50% of its maximum value.

(af) Frequency distortion occurs when the amplifier does not amplify equal signals of different frequencies.

(ag) A differentiating circuit is an RC coupling circuit with a short time constant. A circuit of this type will produce sharp spikes in the output which can be used for controlling stages in a TV receiver.

(ah) An integrating circuit is a circuit with a long time constant that will build a series of pulses up into a single pulse.

Lesson Questions

Be sure to number your Answer Sheet B108.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What determines whether a resonant circuit is a series-resonant or a parallel-resonant circuit?

2. What is the impedance of a series-resonant circuit having the following component values: $R = 25$ ohms, $X_L = 200$ ohms, $X_C = 200$ ohms?

3. What do we mean by the resonant voltage step-up in a series-resonant circuit?

4. In which type of series-resonant circuit will the resonant voltage step-up be the greatest: (a) a high-Q circuit, (b) a low-Q circuit? Why?

5. What part does a parallel-resonant circuit act like at resonance?

6. What is meant by the resonant current step-up in a parallel-resonant circuit?

7. Which one of the following parts does a parallel-resonant circuit act like below the resonant frequency: (a) a resistor (b) a coil (c) a capacitor?

8. Compare the following characteristics of series and parallel circuits at resonance: (a) resistance at resonance, (b) current at resonance.

9. What is the difference between an RC coupling circuit and an RC differentiating circuit?

10. What is the purpose of an integrating circuit?

YOU HAVE AN AIM IN LIFE

When you enrolled as a student member of the National Radio Institute, you took the first step on your road to success and happiness. You now have a goal for yourself -- you have an aim in life -- you are looking forward to the sort of work you like, the sort of income you want, and the respect and admiration of your friends.

Keep your goal in mind. Never forget it for a moment. Of course you will have your moments of discouragement -- we all have. But if you make a thorough search for the cause of your discouragement, you will most likely find that you ate something which did not agree with you, or were kept awake last night by the neighbor's dog.

Whenever you are tempted to neglect your studies, say to yourself: "I have a goal to reach and I'm going to reach it." Think how unhappy you would be if you did not have this goal. There is nothing as pathetic as a rudderless ship or a man without an aim in life.

Here's to success and happiness -- your goal.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-208

ACHIEVEMENT THROUGH ELECTRONICS

TRANSFORMERS, IRON-CORE CHOKES, AND RELAYS

B109

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

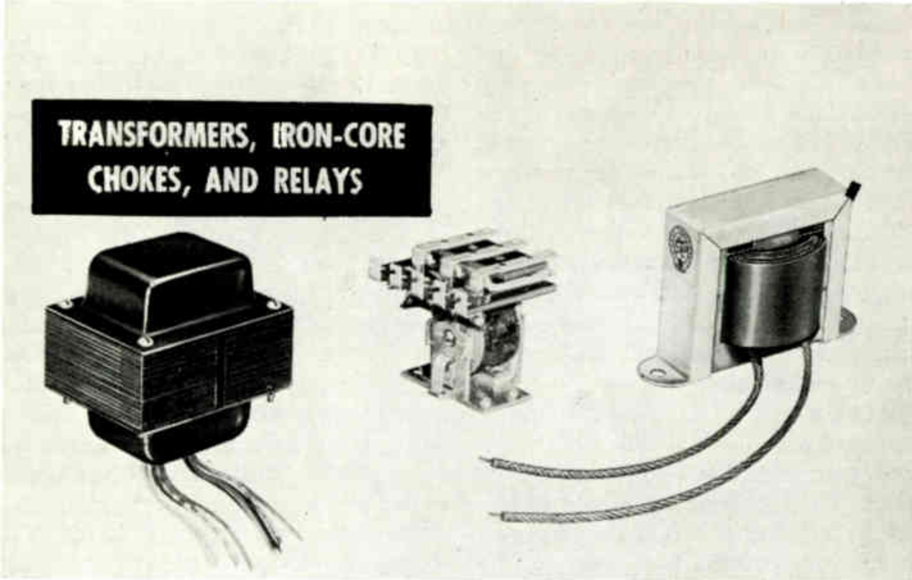
TRANSFORMERS, IRON-CORE CHOKES, AND RELAYS

B109

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

JDL3M1174 1974 EDITION Litho in U.S.A.



In this lesson we will take up three very important parts which all operate on magnetic principles; these are: transformers, iron-core chokes and relays.

Transformers.

You already have an idea of how important transformers are. Transformers are used on ac power lines either to step-up or to step-down the voltage. They are used in electronic equipment to supply the high voltage needed to operate the plates of the various tubes in the equipment and to supply the low voltage to heat the tubes. They are used in transistorized equipment to reduce the power line voltage to the voltage required by the transistors.

Transformers have uses other than supplying electric power. They are used in coupling circuits to transmit an audio signal from one stage to another. They are also used in what are called "matching" circuits to connect a device of one impedance to a device of another impedance. An example of this is the output transformer in a radio receiver, which is used to couple the low-impedance speaker to the higher impedance output tube or transistor that drives the speaker. If it were not for the transformer it would be difficult to get enough power from the tube or transistor to drive the speaker.

Transformers are also used in television receivers to transfer the power from the sweep circuits to the deflection yoke. The deflection yoke is used around the neck of the picture tube to move the electron beam over the face of the picture tube to reproduce the television picture. If it were not for transformers, it would be difficult to get the power needed into the deflection yoke to move the beam over the face of the picture tube.

Iron-Core Chokes.

Iron-core chokes are found most frequently in power-supply equipment. Chokes are used to help smooth the pulsating dc found at the output of a rectifier into pure, ripple-free dc. When chokes are used for this purpose, they are called "filter" chokes because they "filter" the ripple or hum out of the pulsating dc. You will see later that this is possible because the reactance of an iron-core choke is much higher than the dc resistance.

Relays.

There are several types of relays that you are likely to meet in your electronics career. One type of relay is nothing more than a form of automatic switch. The relay can be made either to close or to open the switch when power is applied to it.

There are also other types of relays; for example, some are used to protect circuits. This type is adjusted so that if the current flowing through it exceeds a certain value, the relay will automatically open the circuit, thus protecting the device from an overload.

Another type of relay is called a time-delay relay. This kind of relay is often found in electronic equipment where it is important for the heaters of the various tubes to have time to heat before the high plate voltage is applied to the tube. This type of relay is energized when the equipment is turned on. After a predetermined time, the relay operates, closes the circuit, and applies plate voltage to the tubes in the equipment.

Since the parts we'll study are all magnetic devices, before we look into any of them, we will review what you have already learned about magnetic circuits and learn some additional facts about these circuits.

Magnetic Circuits

As we learned previously when we studied magnetic circuits, we can compare them with electric circuits. The force that drives the flux through a magnetic circuit is the magnetomotive force. This force can be compared to the electromotive force that drives current through an electric circuit. The flux that is driven through the magnetic circuit resembles the current that is driven through an electric curcuit.

The opposition to the flux through the circuit can be compared to resistance in an electrical circuit, and is called reluctance.

MAGNETIC UNITS

There are units set up to measure many of the quantities encountered in magnetic circuits. It is not important for you to remember these units, so do not try to memorize them. We are presenting them here, however, so that you will have seen these terms and will have an idea of what they are when you run into them in the future. After you have completed your NRI course, you will have to keep abreast with new developments. In reading the literature on new developments in the electronics field, it is quite possible that you will run into many of these terms.

Units of Magnetomotive Force.

The magnetomotive force is expressed in terms of ampere-turns. If a current of 1 ampere flows through a coil having one turn, the magnetomotive force developed is 1 ampere-turn. If the coil had two turns, then the magnetomotive force would be 2 ampere-turns, and if a current of 2 amperes flows through a two-turn coil, the magnetomotive force will be 4 ampere-turns.

The ampere-turn is an entirely satisfactory term for use in expressing the magnetomotive force of an electromagnet. However, it is not suitable for use with permanent magnets, and for this reason, another unit of magnetomotive force is used. This unit is the gilbert. The gilbert is slightly smaller than the ampere-turn. To convert ampere-turns to gilberts, you multiply the number of ampere-turns by 1.25.

The magnetomotive force is the total force acting throughout the length of the entire magnetic circuit. Sometimes we want to express the magnetic force in terms of the magnetomotive force per centimeter. (The centimeter is a metric unit of measurement. There are about 2.5 centimeters in an inch.) It may be given as gilberts per centimeter. Then, if the length of the magnetic circuit is six centimeters and the magnetomotive force per centimeter is 10 gilberts, the total magnetomotive force would be 60 gilberts. The magnetomotive force per centimeter is called the magnetic force or the magnetizing force.

Units of Flux.

Previously when we were discussing magnetic flux we simply referred to the number of lines of flux. However, there is a term used for this purpose and it is the maxwell. One line of flux is equal to one maxwell. If you have a hundred flux lines, then the strength of the flux is 100 maxwells. Another unit is the kilomaxwell, which is equal to 1000 maxwells.

Another term that you will encounter is flux density. The flux density tells you how many maxwells or lines of flux pass through a given area. Flux density could be expressed in terms of maxwells-per-square-inch, or it could be expressed in terms of maxwells-per-square centimeter. A flux density of 1 maxwell or one line per square centimeter is known as a gauss. Thus if we say that the flux density is 100 gausses, we mean that there are 100 lines for each square centimeter of a cross-sectional area. If you had a magnet that was 3 × 5 centimeters, the total cross-section of the area would be 15 square centimeters. If the flux density is 100 gausses, then the total number of lines flowing would be 15 times one hundred or 1500 maxwells.

Units of Reluctance.

There is no unit of reluctance. Engineers and technicians are more concerned with the permeability of a material, which you might say is the opposite of reluctance. It is the ability of the material to conduct magnetic flux. It is similar to conductivity in an electric circuit, which is the ability of the material to conduct an electric current.

There is no unit of permeability. The permeability of magnetic materials is rated according to how much better the material conducts magnetic flux than air does. The permeability of air and all other non-magnetic materials is given the numerical value of 1. If the permeability of a magnetic material is 2, a magnetomotive force applied to it will produce twice as many flux lines as the same force applied to air. If we say the permeability of a certain magnetic material is 100, we mean that if the magnetomotive force applied to this material was applied to air, and it produced one line or one maxwell in air, then it would produce 100 flux lines or 100 maxwells in the material.

MAGNETIC SATURATION

One very important thing you should know about magnetic circuits is that they can be saturated. This means that they reach a point where all the possible lines of force exist, and increasing the magnetomotive force applied to the circuit will not produce any further increase in flux. Now let's see how this can happen.

A magnetic material such as iron is actually made up of millions of small molecules. A molecule is a tiny particle made up of a combination of two or more atoms. Each of these molecules is actually a small permanent magnet having a north pole and a south pole. When there is no magnetomotive force applied to the material, the molecules are arranged in a helter-skelter fashion as shown in A of Fig. 1. You will notice that the magnets in this material are pointing in all directions; there is no general organization so that all the north poles point in one direction and all the south poles point in another.

Let's see what happens if we apply a magnetomotive force to this material. This can be done by winding a

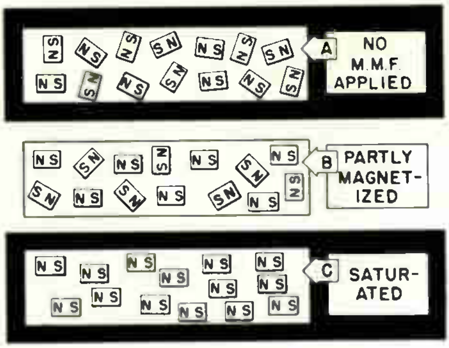


Fig. 1. How the molecules in a material line up as magnetomotive force is applied. Each molecule is like a tiny magnet.

coil around the material and passing a current through the coil. If the current is strong enough to partly magnetize the material, some of the molecules will line up as shown in Fig. 1B. Notice in this figure that there is a general tendency for the north poles to point towards the left and the south poles to point to the right. However, there are a number of molecules that do not follow this general pattern. Some of them are still not lined up.

If we increase the current flowing through the coil or if we increase the magnetomotive force by keeping the current constant and putting more turns on the coil, we will eventually reach a point where all of the molecules are aligned as in C of Fig. 1. Here all the north poles point to the left and all the south poles point to the right. When this situation exists, we say that the material is saturated. This means that any further increase in the magnetomotive force will not produce any further increase in flux. You can see why this is so--all the molecules are already aligned; therefore, it would be impossible to line any more of them up and get more flux. Thus, there would be no point in increasing the magnetomotive force. As a matter of fact, saturation is a condition to be avoided in most cases and if the current is increased beyond the amount needed to align all the molecules, some very undesirable results may occur.

Saturation is sometimes referred to simply as saturation; other times it is referred to as "core saturation" because the magnetic material is usually the core of some device. The magnet may be used as the core of a transformer or a choke. Only magnetic materials can be saturated; an air core cannot be saturated.

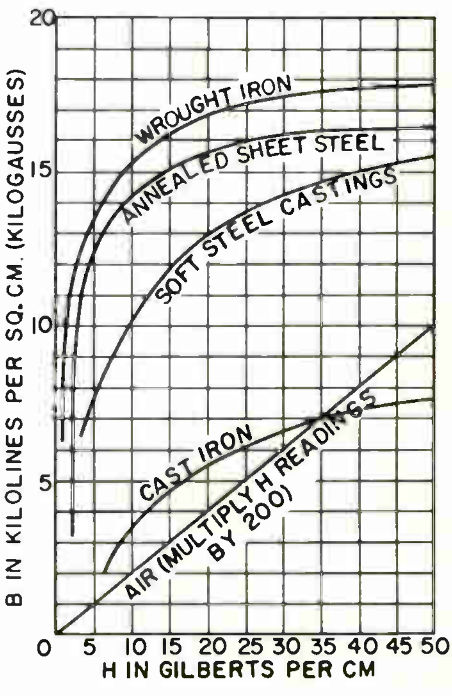


Fig. 2. B-H curves for different magnetic materials.

B-H Curves.

The characteristics of a magnetic material are often represented in the form of a curve called a B-H curve. B-H curves for several different materials are shown in Fig. 2. These curves show the flux density that can be obtained with a given magnetizing force. Notice that as the magnetizing force starts to increase from zero at the left of the graph, the flux density increases quite rapidly at first, then a point is reached where the curve starts to flatten out and it takes a substantial increase in magnetizing force to get even a small increase in flux density. Eventually a point is reached at which the flux density increases no further regardless of how much the magnetizing force is increased. This is the saturation point. However, notice on this graph that the curve for air is a straight line. This simply means that as long as we continue to increase the magnetizing force, the flux density in air will increase. In other words, an air core coil cannot be saturated.

Curves of this type are often used by manufacturers of magnetic materials to describe the characteristics of these materials.

IRON-CORE LOSSES

When we were discussing saturation, we spoke of gradually increasing the current flowing through the coil to increase the magnetomotive force applied to the material. We were discussing a dc current flowing through the coil. Even though we increased the current to show the effects of saturation, the current was still flowing in the same direction. However, we are greatly concerned with the action of coils when alternating current flows through them and the effect the ac has on the magnetic circuit. One of the important things to consider when dealing with alternating current is the losses produced in the iron-core itself.

Hysteresis.

Suppose we apply dc to a coil which in turn produces a certain magnetizing force. As the magnetizing force increases, the flux density increases, as shown by the curve A in Fig. 3, following a curve such as the B-H curves shown in Fig. 2. At zero, the material is not magnetized, and as we increase the magnetizing force, the flux density increases. Now suppose we increase the force only up to point 1 in Fig. 3, then we decide that we will not increase the current any further, but instead

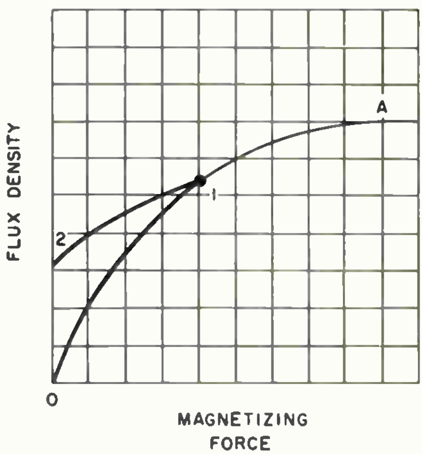


Fig. 3. Curve showing hysteresis loss in a magnetic material. If there were no loss, the curve 1-2 would coincide with the curve 0-1.

gradually decrease the current flowing through the coil, thus reducing the magnetizing force. As the magnetizing force is reduced, the flux density will decrease. However, instead of dropping back down to zero, it will follow the curve shown between points 1 and 2. At point 2, the magnetizing force has been removed entirely, but there is still a certain amount of flux. To get rid of this flux, we must actually reverse the current flowing through the coil. The power that must be applied to bring the flux density back to zero represents a loss due to the inertia of the magnetic circuit. This loss is called the hysteresis loss, pronounced hiss-ter-E-sis.

Any iron-core device operated from ac will waste part of the power applied to it in this way. There is no way that we can eliminate the hysteresis loss altogether, but the amount of power lost will depend upon the material used. By the proper choices of material, the loss can be kept to a minimum. For example, hard steel retains its magnetism and therefore the hysteresis loss in a material of this type would be quite high. On the other hand, soft iron and silicon steel retain very little of the magnetism so the hysteresis loss in the material of this type is much less than in hard steel. For this reason, the iron cores used in most transformers and chokes are made of silicon steel.

Eddy Current Losses.

Another loss that occurs in all iron-core devices is known as eddy current loss.

You will remember that when a varying current flows through a coil, a varying flux is produced in the coil.

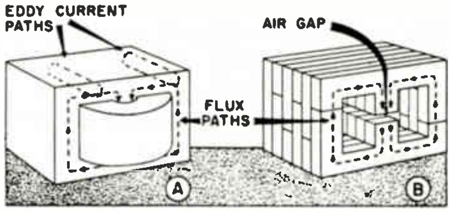


Fig. 4. If the core of a transformer is solid as at A, there are many eddy current paths in it; these can be reduced by making a laminated core as at B.

This flux will cut the turns of the coil and induce a voltage in it. The flux also cuts the turns of any nearby coil and induces a voltage in it, and if the circuit to this coil is complete, this induced voltage will cause a current to flow through it. If a transformer is made like the one shown in Fig. 4A, there are actually many paths in the core. Each path acts like a single-turn coil. We have shown two of these paths where a voltage can be induced which will result in a current flowing. These currents are called eddy currents and represent a loss.

The eddy current loss in a transformer core can be kept at a minimum by making the core out of thin sheets of magnetic material called laminations. The sheets are insulated from each other by a coating of shellac or some other non-conductive material. The sheets are then stacked as shown in Fig. 4B.

Eddy current losses cannot be completely eliminated because even though the core is made of thin sheets, there are still some complete paths present that act like single shorted turns. However, making the core of thin laminations reduces the eddy current losses to a low value.

Both eddy current losses and hysteresis losses vary with the frequency. If the frequency is increased, both losses increase. Thus, although these losses do present some problem at power-line frequencies, they present an even greater problem at audio frequencies, which may extend as high as 15,000 cycles or more. These losses also explain why even laminated iron-core transformers are of no value at radio frequencies. The losses become so high that all of the energy put into the primary of the transformer would be converted into heat due to the eddy current and hysteresis losses. At radio frequencies, magnetic cores are made of finely ground powdered iron mixed with a binder to hold the particles together and insulate them from each other.

Flux Leakage Losses.

Unfortunately not all the flux produced by the magnetomotive force applied to a magnetic circuit will flow through the iron core of a device such as a transformer. Part of the flux will escape and travel through the air surrounding the core. This flux serves no useful purpose since it leaks out of the core and these flux lines do not cut the turns of the secondary winding. This loss is referred to as flux leakage loss.

In a device such as a transformer, if the magnetic material used in the core approaches the saturation point, the flux leakage losses become quite high. Therefore, to keep this type of loss as low as possible, transformers are usually designed to operate well below the saturation point. However, even then it is impossible to eliminate this loss completely, because a certain amount of the flux produced by the primary will travel in a path other than through the core.

Flux leakage is important not only because it represents a loss in the transformer, but also because the escaping flux lines may cut through some nearby part and induce a voltage in it. Thus energy in the transformer can be unintentionally fed into some other part. The amount of energy fed back may be high enough to upset the performance of the equipment.

Flux leakage can also present a problem in television receivers. Flux leaking from a transformer can deflect the electron beam in a picture tube and cause distortion in the picture. In color TV receivers flux leakage can actually cause the color picture to break up into three separate pictures of different colors or cause color fringing where objects are outlined in one or more colors.

In designing electronic equipment using transformers, engineers must consider the possibility of these undesired effects and try to keep transformer leakage fields away from the picture tube in TV receivers or other parts in the electronic equipment that could pick up interference from the field. In equipment where several transformers are used, they try to place the transformers so that there will be a minimum of interaction between them.

SUMMARY

In this section, you reviewed the facts you previously learned about magnetic circuits. You also learned some new terms. A unit of magnetomotive force that you will encounter is the gilbert and a unit of flux is the maxwell. You learned that the magnetomotive force is the force supplied throughout the entire magnetic circuit. Sometimes the force is expressed in terms of the magnetomotive force per unit length and is called the magnetizing force.

The amount of flux produced is sometimes expressed in terms of so many lines of maxwells for a given area. If the unit of cross-section area used is the centimeter, and you have 1 maxwell per square centimeter, we say the flux density is 1 gauss.

It is important for you to remember that a magnetic material can be saturated and that when the saturation point is reached, increasing the magnetomotive force will result in no further increase in flux.

The losses encountered in iron cores are important. The three most important losses are hysteresis loss, eddy current loss, and flux leakage loss. All these losses represent some energy that is being put into the primary of the transformer for which we get nothing out of the secondary. In the next section we will study transformers and you'll see the importance of these losses.

SELF-TEST QUESTIONS

(a) What is the gilbert?

(b) What is the unit of magnetic flux?

(c) What is meant by flux density?

(d) If the cross section area of a magnet is ten square centimeters and the flux density is ten gausses, what will the total number of flux lines flowing be?

(e) What is meant by magnetic saturation?

(f) Is magnetic saturation a desirable condition?

(g) How strong a magnetomotive force is required to produce saturation in air?

(h) What is a hysteresis loss?

(i) What are eddy current losses?

(j) How are hysteresis losses kept at a minimum?

(k) How are eddy current losses kept to a minimum?

(l) What are flux leakage losses?

Iron-Core Power Transformers

The ability of a varying magnetic field to induce a voltage in any conductor with which it links makes it possible to transfer power from one circuit to another without direct wiring connections. The device used for this purpose is called a transformer. In its simplest form a transformer is nothing other than two separate coils of wire wound on a common core or wound in such a way that the coils of wire are placed near each other so that the magnetic lines produced by one will cut the other.

A power transformer is wound on an iron-core. The iron-core is usually shaped like the core shown in Fig. 5A. The coils are wound on the center leg of the core as shown in Fig. 5B, one inside the other. The schematic symbol used to identify an iron-core transformer is shown in Fig. 5C.

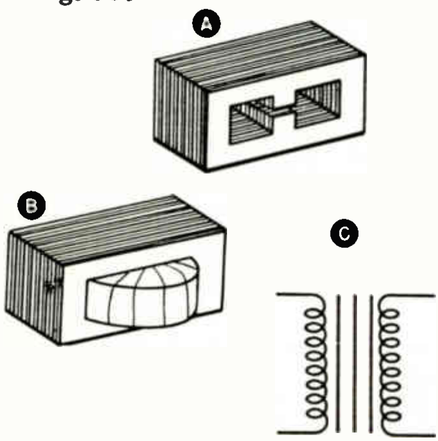


Fig. 5. The core of a power transformer is shown at A. The coils are wound on the center leg as at B. The schematic symbol for an iron-core transformer is shown at C.

Transformers are extremely important to the electronics technician, so let's learn more about them.

POWER LOSSES

In the preceding section of this lesson you learned about some of the losses that take place in magnetic cores. These losses, which are the hysteresis, eddy current, and flux leakage losses, are called "core" losses because they are characteristic of the magnetic core which is used in the transformer. However, there are also other losses in transformers.

The coils making up the transformer are wound of copper wire. Copper has a very low resistance, but nevertheless it does have some, so the current flowing through the copper wire will encounter opposition. This means power will be lost or used in forcing the current through the coil.

You know that the power is equal to voltage multiplied by the current; in other words:

$$P = E \times I$$

But from Ohm's Law we know that:

$$E = I \times R$$

Therefore, in the power formula we can substitute I × R for E, and get

$$P = I \times R \times I$$

which is usually written in the form

$$P = I^2R$$

The power lost as a result of current flowing through copper wire in a transformer will be equal to the current squared times the resistance of the wire. This loss is usually referred to as the "I squared R" loss.

Thus, in a transformer we have two groups of losses: the core losses and the copper losses. These losses appear in the form of heat. When a transformer is put into operation, it starts to heat up. The transformer will continue to get hotter and hotter until eventually a state of balance is reached where further heat generated by the transformer can be carried away by the air surrounding the transformer and by the metal chassis on which the transformer is mounted. When this balance is reached, the temperature of the transformer will stop rising, and the transformer will not get any hotter.

The amount of heat that a transformer can dissipate in this way depends upon its size. Since a large transformer will have a much greater air circulation and will be mounted on a larger area of the chassis, it can get rid of more heat than a small transformer. Therefore a large transformer is capable of handling a larger amount of power than a small transformer. As a matter of fact, the size of the transformer is usually a pretty good indication of the amount of power that it can handle safely. After you have worked on electronic equipment for a while, you'll learn to recognize from the size of a transformer approximately how much power it can handle.

Transformers can be overloaded if too much power is taken from them. When a transformer begins to overheat it is an indication that it is being overloaded. Usually when a transformer is overheated in this way you will notice a dark colored sealing compound or wax leaking out of the transformer, and often you can smell the varnish and insulation burning.

An overload of this type may be due to a defect in the transformer itself or it may be due to a defect somewhere in the equipment that is pulling excessive current from the transformer. If the transformer itself is defective, the trouble is usually that two or more turns of the transformer have touched each other so that a short circuit exists and current can simply flow around inside the transformer. When this happens there is nothing you can do except replace the transformer. However, if the defect is in the equipment rather than the transformer, and if you find it and eliminate the defect before operating the equipment any more, the chances are that the transformer will once again give satisfactory service.

If the transformer is operated with an overload, it will eventually get so hot that the insulation on the copper wire and on the terminal used to insulate one winding of the transformer from another will become overheated. Usually paper is used to insulate the various windings on a transformer. If this paper becomes too hot, it becomes charred so that it is no longer a good insulator, and the windings on the transformer will short together. When this happens, the transformer is no longer usable because it will draw more and more current and it will get hotter and hotter and eventually it will blow a fuse or the copper wire on one of the windings will melt so that the winding opens.

Transformers are designed to operate on a specific frequency. In other words, a transformer designed to operate on a 60-cycle power line will operate best only on a power line of that frequency. If you operate a 60-cycle power transformer on a 25-cycle power line or accidentally plug it into a dc power line, the transformer will burn out. A 60-cycle transformer cannot be operated on a 25-cycle power line. No transformer can be operated from dc. A 25-cycle transformer will operate on a 60-cycle power line, but 25-cycle transformers are much larger and much more costly to manufacture than 60-cycle transformers, and therefore it would be uneconomical to design a transformer for 25-cycle power and then use it on a 60-cycle power line.

In spite of the losses we have discussed, a transformer is one of the most efficient devices you will ever find. Large transformers such as those used by the power company achieve a very high efficiency, usually from 98% to 99%. Smaller transformers such as you will find in electronic equipment usually operate at an efficiency of somewhere between 95% and 98%. As far as the technician is concerned, this high efficiency means that in many cases you can ignore the transformer losses. You can consider the power output of the transformer as being equal to the power input; the difference will be only a small percent of the total power, and in evaluating the performance of the transformer, very little error will be introduced.

TURNS RATIO

The turns ratio of a transformer is the ratio of the number of turns on one winding of the transformer to the number of turns on the other winding. For convenience we usually identify the two windings on a transformer as the primary winding and the secondary winding or, more simply, as the primary and the secondary. The primary is the winding to which we apply the input power. The secondary winding is the winding from which we take power. If the primary winding of the transformer has 500 turns, and the secondary winding has 100 turns, we say that the transformer has a turns ratio of 5 to 1. This is often written 5:1. If the primary of the transformer has 100 turns, and the secondary 500 turns, then we say that the transformer has a turns ratio of 1 to 5 (1:5).

The ratio of the secondary voltage to the primary voltage will depend upon the turns ratio. If the secondary winding has five times as many turns as the primary, we can expect to get five times the voltage from the secondary that we put into the primary. Similarly, if the secondary has only half as many turns as the primary, then we can expect to get half the voltage across the secondary that we put into the primary. If we get more voltage out of the secondary than we put into the primary, the transformer is called a "step-up" transformer, and if we get a lower voltage out of the secondary than we put into the primary, then the transformer is called a "step-down" transformer.

In expressing the turns ratio of a transformer, manufacturers and technicians do not always give it as the ratio of the primary turns to the secondary turns. If a transformer has 100 turns on the primary and 200 turns on the secondary, the turns ratio is 1:2. However, sometimes this turns ratio is given as 2:1, step-up. This tells you the transformer is a step-up transformer which means there are more turns on the secondary than on the primary and the secondary has twice as many turns as the primary.

Power Consumption.

One of the things that makes a transformer so useful is that it is basically a self-regulating device. By this we mean it takes no more power from the power line to which it is connected than is needed to supply the power demanded from the secondary. In other words, if we connect a load across the transformer secondary and this load consumes 50 watts, then the power drawn from the power line by the primary will be 50 watts. Similarly, if we connect a 100-watt load across the secondary, then the power drawn from the power line by the primary will be approximately 100 watts. The primary will draw approximately the power required in order to furnish the demands of the secondary. We say approximately, because there are losses in the transformer itself, and the primary will draw the power needed to supply the secondary power plus the power needed to supply the losses within the transformer itself.

Consider what happens when the primary of the transformer is connected across the power line, but there is no load connected across the secondary. Under these circumstances, the secondary power is zero. Therefore, the primary does not have to supply any power to the secondary, and the power it will consume for this purpose will also be zero. The only power that the primary will consume from the power line will be the power to make up the core losses and a very small copper loss. Thus, when the primary of a transformer is connected across the voltage source and there is no load connected to the secondary, the transformer draws very little power from the voltage source. Under these circumstances, the transformer is very inefficient because all the power it is consuming is being wasted. However, as we load the secondary, these losses remain almost constant and as the secondary begins to use power, the primary power increases until when the transformer is being operated at its rated power, its efficiency reaches a very high value.

Because the primary power depends upon the secondary power, the actual current that will flow through the primary of a transformer will depend upon the load connected to the secondary. Ignoring the core and copper losses, if we have a 10 to 1 step-down transformer connected across a 100-volt power line, the voltage available at the secondary will be 10 volts. The turns ratio determines the ratio of the primary to secondary voltage. The ratio of the current flowing in the primary to the current flowing in the secondary also depends on the turns ratio, but it works in the opposite way. If we connect a load across the secondary that draws a current of 1 amp, then the power taken from the secondary will be 10 volts times 1 amp, which equals 10 watts. To supply 10 watts, the primary, since the voltage across it is 100 volts, needs only 1/10 of an amp and therefore this is the current that it will draw from the power line. If we increase the load across the secondary and pull 100 watts from the secondary, then the current flowing in the secondary must be 10 amps, because with a voltage of 10 volts, it will take a current of 10 amps to supply 100 watts. Under these circumstances, the primary again will pull the power needed from the power line to supply this 100 watts. This means that the primary current will be 1 amp.

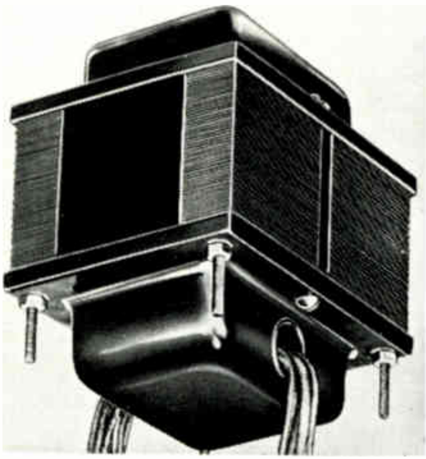
Now notice what our situation is. Here we have a step-down transformer with a 10-to-1 turns ratio. This transformer steps the voltage from 100 volts down to 10 volts. However, the current acts in the opposite way. The secondary current is higher than the primary current. The secondary current will actually be 10 times the primary current if we ignore the transformer losses. Therefore in a step-down transformer, the current is stepped up and similarly in a step-up transformer where the voltage is stepped up, the current is stepped down.

Since the current drawn from the primary will vary as the loading on the secondary of the transformer varies, this means that the impedance of the primary winding must vary. If the impedance of the primary winding remained constant, then the current flowing through the transformer primary would be constant for any given primary voltage. However, since the current does vary, then the impedance must vary. This it does, in fact, and the actual impedance of the primary depends upon the impedance connected across the secondary. This is caused by the fact that when the impedance across the secondary varies, the current as well as the power demanded from the secondary will vary. The primary current and the impedance will also vary in turn.

TYPICAL POWER TRANSFORMERS

Many pieces of electronic equipment that you will service will use power transformers. Power transformers serve a number of useful purposes. First, by using a power transformer it is possible to have available a number of different operating voltages other than the single voltage available directly from the power line. Furthermore, the power transformer isolates the equipment from the power line. This is a big advantage because one side of most power lines is grounded. If you accidentally come in contact with any grounded object and some of the circuits of the electronic equipment or the metal chassis of the equipment at the same time, it is possible to get a severe shock from electronic equipment that does not use a power transformer.

The power transformer found in modern electronic equipment has a primary winding and one or more



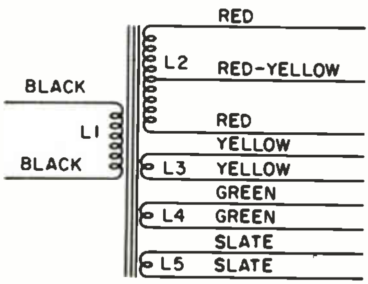


Fig. 6. A typical power transformer and its schematic symbol.

secondary windings. A photo of a typical power transformer and the schematic symbol used to represent it are shown in Fig. 6.

The winding marked L1 is the primary winding. It is usually operated from a 115-volt, ac power line.

The secondary winding marked L2 is a high-voltage secondary. Notice that this winding is center tapped. This type of winding is used with a full-wave rectifier.

A full-wave rectifier is a rectifier that rectifies both halves of each cycle. This winding is a stepup winding and is used to provide a voltage somewhat higher than that available from the power line to operate the plates of the various tubes in electronic equipment. Voltages of 250 to 350 volts are found in most circuits of modern radio and TV receivers; much higher voltages are found in transmitting equipment and in other pieces of industrial electronic equipment.

The winding marked L3 is a stepdown winding. This winding is used to provide the filament voltage to operate the filament of the rectifier tube, which is used to change the ac to pulsating dc. The windings marked L4 and L5 are also low-voltage windings. These windings are used to provide the heater voltage required by the various tubes in the equipment. Some transformers have two low-voltage secondary windings like L4 and L5 on this transformer, but others have only one low-voltage secondary to heat the various tubes.

Notice that the number of turns used in the schematic symbol gives some indication of whether the windings are step-up or step-down windings. L2 has more turns than L1, and here you can expect the voltage from L2 to be higher than the primary voltage applied to L1. The windings L3, L4, and L5 have fewer turns than L1, indicating that their voltage is less than that of the primary winding. However, the schematic is not intended to show the exact number of turns or the turns ratio.

The colors have been labeled on the various leads from this transformer. This is a standard color

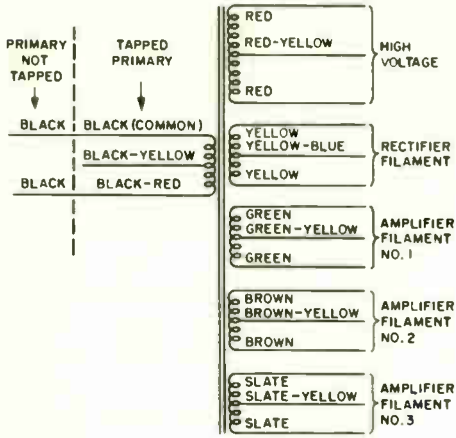


Fig. 7. The standard EIA (Electronics Industries Association) color code for power transformers.

code used to identify transformer leads. The complete color code is shown in Fig. 7. However, do not expect all transformers to follow this standard color code. Some manufacturers use a color code of their own. Therefore, even when you are using the color code to identify the leads on a transformer, you should pay some attention to how the transformer is connected into the circuit, the size of the various wires, etc., to be sure that the manufacturer has followed the standard code.

SUMMARY

We have not tried to cover all the facts about iron-core power transformers in this section of this lesson. Transformers are a subject all by themselves. Some engineers spend their whole careers designing different types of transformers. However, the information covered in this section will enable you as a technician to understand enough about transformers to know how they operate. You will learn still more about audio transformers in this lesson.

The important thing to remember from this section is that there are a number of different types of losses in transformers and that these losses cause the transformer to heat. Losses can be divided into core losses which consist of eddy current, hysteresis, and flux leakage losses, and copper losses which are called the "I squared R" losses.

Normally a transformer heats up as it is used, until it reaches a point at which it does not get any hotter. If the transformer in a piece of electronic equipment continues to get hotter and hotter and you can smell the insulation burning, it is an indication either that it is being overloaded or that there is a short either in the transformer or in the rest of the equipment.

A step-up transformer is a transformer where a higher voltage is obtained from the secondary than is put into the primary, and a step-down transformer is a transformer where a lower voltage is obtained from the secondary than is put into the primary. The power output from the secondary of a transformer is approximately equal to the power input of the primary. Thus in a step-up transformer where the secondary voltage is higher than the primary voltage, the primary current must be higher than the secondary current. Conversely, in a step-down transformer where the secondary voltage is lower than the primary voltage, the primary current will be lower than the secondary current.

Power transformers used in electronic equipment usually have several secondary windings. One secondary winding is used to provide the high voltage needed to operate the plates of the various tubes. The other secondary windings usually supply the heater voltage required by the rectifier and the other tubes in the equipment. Most transformers use a standard color code which can be used to identify the various transformer leads.

SELF-TEST QUESTIONS

(m) What is a transformer?

(n) What are the two types of losses encountered in a transformer?

(o) What happens to the power lost in a transformer?

(p) Is a transformer generally considered an efficient device?

(q) Can a power transformer designed for operation on a 60-cycle power line be used on a 25-cycle power line?

(r) What do we mean by a step-up transformer?

(s) If you have a step-down power transformer with a turns ratio of 3:1, and the current being drawn from the secondary is 3 amps, what will the primary current be?

(t) The load connected across the secondary of a power transformer consumes 230 watts. If the primary of the transformer is operated from a 115-volt power line, what will the primary current be if we consider the transformer efficiency as 100%?

(u) A power transformer has a turns ratio of 2:1. The transformer is operated from a 120-volt power line, and the device connected to the secondary winding draws a current of 4 amps. What will the primary current be?

Transformers for Specific Application

As an electronics technician, there are a number of different types of transformers that you will encounter. We are not going to try to cover all of them here, but we will discuss a few of the more common types. The material in this section of the lesson is used simply to introduce you to these special types. Later we will go into the transformers in more detail when we study the applications in which they are used.

AUDIO TRANSFORMERS

In the early days of radio, audio transformers were used as coupling devices between the various stages of the receiver. By using a step-up transformer, the strength of the audio signal could be increased in the transformer itself. This was a big help, because the tubes used in those days did not have a great deal of gain. However, modern tubes have high gain and audio transformers introduce frequency distortion, so they are no longer used with tubes for this purpose.

Impedance Matching.

You will remember that to transfer maximum power from a dc generator to a load, the load resistance must be equal to the generator resistance. In ac circuits, to transfer maximum power from a generator to a load, the load resistance must be equal to the generator impedance. To transfer maximum power from one stage in an amplifier to the following stage, the impedances must be matched. Transformers are frequently used for this purpose, particularly in transistorized equipment. For example, the output circuit of a transistor may have a much higher impedance than the input circuit of the following transistor. To transfer the power from the one transistor to the second one, an impedance-matching transformer is used.

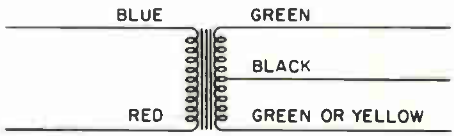


Fig. 8. An audio transformer with a tapped secondary. The colors are the standard EIA color code.

Another application in which a transformer may be used is where a single tube or transistor is used to drive two tubes or two transistors. Then an audio transformer with a tapped secondary such as shown in Fig. 8 is used. With this type of transformer, the center tap winding is either connected to ground or used to supply bias to the stage. The two tubes or transistors are driven with signals of the opposite polarity. In other words, when the signal on the top green lead in Fig. 8 is swinging positive, the signal on the lower lead that is green or yellow will be swinging negative. Thus when the current in one of the stages is increasing it will be decreasing in the other and vice versa. This type of stage is called a push-pull stage and is frequently used in the output stage of high fidelity equipment in order to obtain a high power output with low distortion. We will cover this in detail later when you study tubes and transistors.

Another place where transformers are widely used in audio circuits is in the output between the last amplifier stage and the loudspeaker. The output impedance of the power output tube or the power output transistor is generally much higher than the speaker voice coil impedance. A transformer, which is called the output transformer, is used to match the output stage to the loudspeaker in order to transfer maximum power from the output stage to the speaker. A transformer of this type will have a step-down turns ratio in order to match the higher impedance of the output stage to the lower impedance of the speaker.

You saw in an earlier lesson that to transfer maximum dc power from a generator to a load, that the generator and load resistances had to be equal. Now let's see how a transformer can be used in an ac circuit to match the load impedance to the generator impedance, so that the generator will deliver maximum power to a load that has an impedance different from the generator impedance.

In Fig. 9, we have shown how the power supplied to the load by a generator that has a no load voltage of 100 volts and an internal impedance of 50 ohms varies as different load resistors are connected across it. You will notice that when the 50-ohm load is connected across the generator, we obtain 50 watts across the load. If the load impedance is reduced below this value, then the power transferred from the generator to the load drops off.

Similarly, if the load impedance is increased above 50 ohms, the power drops off.

With a 50-ohm load connected across the generator, the total circuit resistance is 100 ohms: the 50-ohm load, plus the 50-ohm generator impedance. The 100 volts generated will be divided with 50 volts being dropped in the generator and 50 volts in the load. The current flowing in the circuit will be 1 amp.

LOAO RESISTANCE IN OHMS	POWER IN WATTS
200	32
150	37.5
100	44.4
50	50
25	44,4

Fig. 9. Power supplied by a 100-volt generator with an internal impedance of 50 ohms for various values of load resistance.

If we remove the 50-ohm load resistor, and connect some other device in its place that will draw a current of 1-amp from the generator, then there will be 50 volts across this device and the power supplied to it will be 50 watts, which we know is the maximum power that can be taken from the generator. A transformer can be used for this purpose providing it has the proper turns ratio, and providing the correct load is connected across the secondary of the transformer.

The usual problem is to have the generator and the load and then have to select the transformer. Let's assume that we have a 2-ohm load resistor we want to connect across the generator and see how a transformer can be used to get 50 watts from the generator to the load.

We know that the maximum power that can be transferred from the generator to the load is 50 watts. Now let's find out what the voltage across the resistor and the current through the resistor must be in order to get 50 watts into the resistor. Remember the formula:

$$P = I^2 \times R$$

We know that P = 50 watts, and R = 2 ohms; therefore we can find the value of I^2 in this manner:

$$I^2 = 50 \div 2 = 25$$

This means that the current squared is equal to 25. We know that the square root of 25 is 5, because 5 × 5 = 25. Therefore the current that must flow through the resistor is 5 amps. In order to figure how much voltage we need to get a current of 5 amps to flow through a 2-ohm resistor, we use the formula: E = I × R. This gives us 2 × 5, which equals 10 volts.

So far we have found that to get 50 watts into a 2-ohm load resistor, we must have a voltage of 10 volts across it. When we have 10 volts across the resistor, a current of 5 amperes will flow through it, and the power supplied to the resistor will be 50 watts. However, we still have the problem of getting the 10 volts across the resistor.

The transformer presents an easy method of doing this. We know that when we had maximum power transfer from the generator to the load, we had 50 volts across the load. However, in the problem we have set up, we want only 10 volts across the load. The way to satisfy both of these conditions is to connect a transformer across the generator as shown in Fig. 10. If we use a transformer with a turns ratio of 5 to 1, we will be able to satisfy these conditions.

You will remember that the transformer is a self-regulating device. The primary will take the power from the source needed to supply the power demanded by the secondary. Under these circumstances when the generator, transformer, and load are connected as shown in Fig. 10, the primary current that

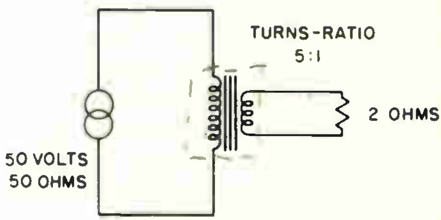


Fig. 10. A transformer used to match a 2-ohm resistor to a generator with an internal resistance of 50 ohms.

will flow will be 1 amp. This means that there will be 1 amp through the primary of the transformer, and therefore the power consumed by it will be 50 watts. The 50 volts applied to the primary will be stepped down by the transformer to give 10 volts across the secondary. With the 2-ohm load connected across the secondary, the secondary current will be 5 amps, and the power consumed by the secondary will be 50 watts. Since the transformer is a step-down transformer, the current is stepped up. The primary current of 1 amp is stepped up to a current of 5 amps in the secondary.

Under these circumstances, as far as the generator is concerned, it works exactly as it would if a 50-ohm resistor were connected across it. The transformer has matched the two-ohm resistor to the generator. Technicians say that the generator "looks into" the load and it looks like a 50-ohm load connected across it. Because the transformer matches the load impedance to the generator to provide the generator with the impedance required to transfer maximum power from the generator to the load, the transformer is called an impedance-matching transformer.

Turns Ratio.

Notice the ratio of the two impedances we matched. The generator impedance was 50 ohms, the load 2 ohms. The ratio of these two impedances is 50 to 2 or 25 to 1. The turns ratio of the transformer, however, was 5 to 1. But 5 is the square root of 25. Thus the relationship between the impedances to be matched and the turns ratio of the transformer needed to match the impedance is:

$$\frac{N1}{N2} = \sqrt{\frac{Z1}{Z2}}$$

where Z1 is the generator impedance, Z2 the load impedance, N1 the number of turns on the primary of the transformer, and N2 the number of turns on the secondary. Thus, N1/N2 is the turns ratio.

AUTOTRANSFORMERS

The schematic of another type of transformer is shown in Fig. 11. As you can see from the schematic, this transformer consists of a single winding with a tap. It is called an autotransformer because it has only one winding. (Auto is the Greek word for "self".) In the transformer shown in Fig. 11A, the secondary voltage is higher than the primary voltage because there are more turns on the secondary than there are on the primary. In the transformer shown in Fig. 11B, the secondary voltage is lower than the primary voltage because the primary has more turns than the secondary.

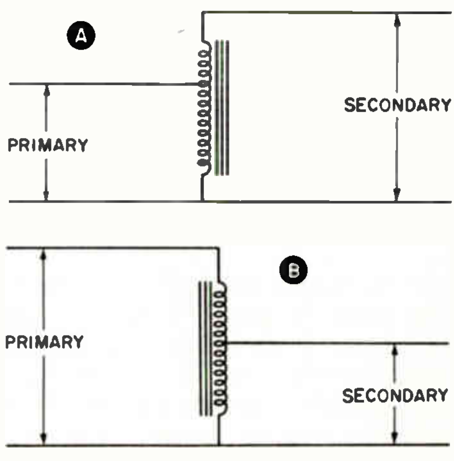


Fig. 11. Schematic of an autotransformer.

The autotransformer is different from a conventional transformer in that the primary and secondary windings are not insulated from each other. The voltage applied to the primary will set up a magnetic field, and this magnetic field will cut all the turns of the secondary, including those which may be a part of the primary, and induce a voltage in them. This voltage will appear across the secondary terminals of the transformer and if we connect a load to these terminals, current will flow through the load.

Autotransformers have the disadvantage that the primary and secondary windings are not completely isolated from each other electrically. In other words, there is an electrical connection between the primary and secondary windings. As a matter of fact, one lead connects directly to both primary and secondary windings as you can see from the diagrams in Fig. 11. However, they have the advantage over the transformer with two separate windings in that they are more economical to manufacture and therefore are frequently used in modern electronic equipment where keeping the cost as low as possible is of major importance.

RF TRANSFORMERS

Another transformer that you will encounter frequently is the rf radio frequency transformer. This type of transformer, because of the high frequencies at which it operates, is either an air-core or a powdered iron-core transformer. A typical rf transformer designed for use at standard broadcast band frequencies is shown in Fig. 12.

RF transformers are used in the stages of a radio or TV receiver that are designed to amplify the received signal frequency. In other words, in the case of a broadcast band receiver, the signal is picked up and amplified by one or more rf stages before it is processed in any way. RF transformers are used between stages of this type.

In most cases, the secondary of the transformer, at least, is used in conjunction with a variable capacitor to form a resonant circuit. The rf transformer acts very much like the step-up transformer, because if the secondary is tuned to resonance, the circuit forms a series-resonant circuit and there will be a resonant voltage step-up across the secondary of the transformer. This situation exists even though the primary of the transformer may have the same number of turns as the secondary. Here the step-up in voltage is being obtained because of the action of the resonant circuit, rather than the action of the transformer.

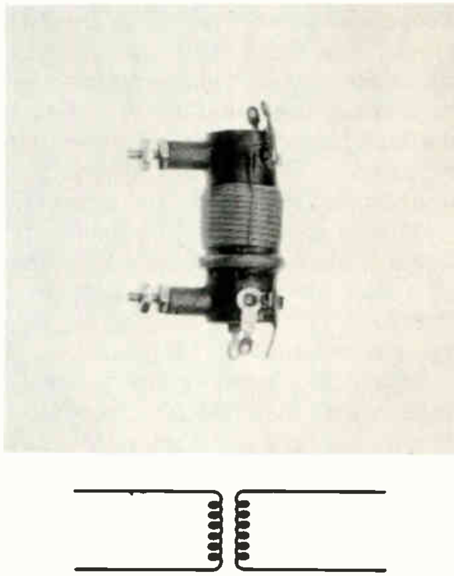


Fig. 12. A typical rf transformer and its schematic symbol.

In TV receivers where the rf stages must operate on a much higher frequency, an rf transformer may consist of only one or two turns of wire on the primary winding and the same number of turns on the secondary. If one winding of the transformer is tuned to resonance, it is usually tuned by stray circuit capacitances and capacitances in the tubes or transistors used in the stage rather than by a separate variable capacitor.

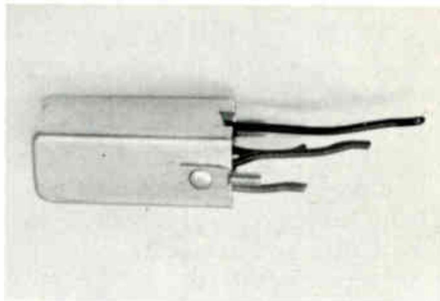
RF transformers are sometimes called rf coils, but are actually transformers. Their operation is similar to that of the iron-core transformers you studied earlier in this lesson.

I-F TRANSFORMERS

Modern radio and television receivers use what is called a superheterodyne circuit. In this type of circuit the incoming signal is fed to a stage called a mixer stage where it is mixed with a locally generated signal. The output of the mixer stage produces a new signal frequency which is called the intermediate frequency. We abbreviate this i-f and call the stages used to amplify this signal i-f stages. Between the various i-f stages we use transformers called i-f transformers.

A typical i-f transformer and its schematic diagram are illustrated in Fig. 13. The primary winding of the transformer and its capacitor form a parallel-resonant circuit, and the secondary winding and its capacitor form a series-resonant circuit. Both the primary and the secondary windings are tuned to the same frequency and usually the two windings have exactly the same number of turns on them. However, again because the secondary is a series-resonant circuit, there is a resonant voltage step-up and the voltage across the secondary winding will be higher than the voltage across the primary winding.

Even though the primary and secondary windings of an i-f transformer are tuned to resonance at a specific frequency, they are de-



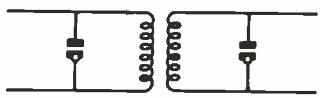


Fig. 13. A typical i-f transformer and the schematic symbol for it.

signed so that they have a certain bandwidth. By this we mean that instead of passing only one specific frequency, the transformer will pass a band of frequencies. The actual bandwidth depends upon the design of the transformer and the frequency at which it operates. The bandwidth of a transformer can be shown by means of a response curve such as shown in Fig. 14. Here you will see that the resonant frequency is 455 kc, but there is very little differ-

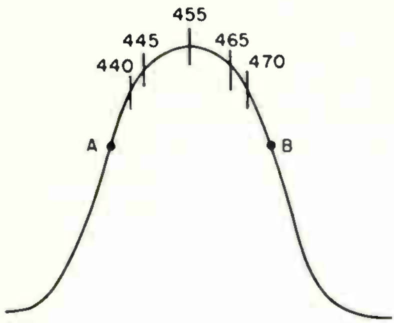


Fig. 14. An i-f transformer response curve.

ence in the response of the transformer either 10 kc below or 10 kc above this frequency. Under these circumstances, the transformer would have a bandwidth of at least 20 kc. In other words, it will pass signals having a frequency from 445 kc to 465 kc satisfactorily.

As a matter of fact, we can go a little further down the curve and see that there is not too much difference between 440 kc and 455 kc. You can see that deciding the bandwidth is a rather arbitrary thing. Engineers have set up as a standard the point where the response falls to .707 of the response at the resonant frequency. This is the half-power point you have already studied. The bandwidth of the i-f coil is the frequency between the .707 point on the low side of the curve and the .707 point on the high frequency side of the curve. These points have been marked A and B on the response curve shown in Fig. 14.

They are called the "3-db down" points. The abbreviation "db" means decibel. It is a means of expressing power ratio. The 3-db down points, A and B in Fig. 14, are within 3-db of the resonant frequency.

We do not expect you to understand the decibel at this time; we will go into it in more detail eventually. However, keep in mind the expression "3-db down", because this expression is frequently used by technicians and engineers. It is an idea that is somewhat difficult to grasp, but one of the first steps in seeing what is meant is becoming familiar with the term. Now that you have been introduced to it, the next time you will see it, it will not seem quite so strange to you.

When you look at Fig. 13 you might think that the primary and secondary windings are placed so far apart that there would be very little coupling between the two of them. However, there actually is considerable coupling between the two windings. The spacing of the transformer windings is adjusted to give exactly the desired amount of coupling. If the coils are placed too close together, then the curve shown in Fig. 14 tends to flatten out and have a dip in the center around a 455 kc point.

In TV i-f transformers, the coils are placed much closer together; in fact, often one is wound directly on top of the other. This is done in order to provide what is called very tight coupling and to spread out the response curve to pass a wide band of frequencies. You will see later that in television we must be able to pass a wide band of frequencies; otherwise, part of the picture information in black and white transmissions or part of the color information in color broadcasts will be lost.

TV i-f transformers do not have

a separate capacitor across the primary and secondary windings like the one shown in Fig. 13. There is enough capacity in the circuit and in the output and input in the various stages to provide the capacity required in order to bring the windings on the transformer to resonance.

SUMMARY

In this section of the lesson, you have been introduced to a number of new types of transformers.

You have learned that transformers are used between the audio stages in some pieces of equipment. Many of these transformers are step-up transformers so that the voltage across the secondary may be two or three times the voltage across the primary. However, in some high-power audio equipment, step-down transformers are used between audio stages. Transformers with tapped secondaries may be used when one tube must drive two tubes.

Transformers are used as impedance-matching devices. In order to get maximum power from a generator to a load, the load impedance must be equal to the generator impedance. This situation can be met by using a transformer with a suitable turns ratio to match the load impedance to the generator impedance so that the generator operates as if it were working into a load equal to its own internal impedance. Impedance-matching transformers are used between the output stage of a radio or TV receiver and the loudspeaker. Because these transformers are used at the output of the receiver they are usually called "output" transformers.

You have learned that the autotransformer is a single-winding transformer. Part of the transformer winding serves as both the primary and secondary. The autotransformer is frequently used to step up the line voltage if it is somewhat lower than normal.

We have also mentioned rf transformers. Although we did not go into a great deal of detail about them, you should recognize an rf transformer the next time you see one. An rf transformer has two windings, a primary and a secondary winding, and it operates in very much the same way as an iron-core transformer.

I-F transformers have both a tuned primary and a tuned secondary. The primary is a parallel-resonant circuit. The secondary is a series-resonant circuit. An i-f transformer has a certain bandwidth; this means that the transformer will pass frequencies above and below the frequency to which it is resonant. The bandwidth of a transformer is defined as the frequency difference between a point on the low side of the response curve and a point on the high side of the response curve at which the output from the transformer is .707 of the output at resonance. These points are called "3-db down" points.

SELF-TEST QUESTIONS

(v) Are step-up type audio transformers used in modern electronic equipment?
(w) What is the major use of audio transformers in electronic equipment?
(x) Why is it important that the

output stage in a radio receiver or a TV receiver be matched to the speaker voice coil?

(y) What type of audio transformer is used to drive a push-pull stage from a single driver stage?

(z) What must the turns ratio of an output transformer be to match a speaker with a 10-ohm voice coil to an output stage that has an output impedance of 1000 ohms?

(aa) What is an autotransformer?

(ab) Is an autotransformer a step-up transformer or is it a step-down transformer?

(ac) What is the disadvantage of an autotransformer?

(ad) What is an rf transformer?

(ae) What type of core would you expect to find in an rf transformer?

(af) In a typical i-f transformer used in a broadcast receiver, the primary and secondary windings will have the same number of turns. However, in spite of this, the voltage across the secondary winding will be higher than the voltage across the primary winding of the transformer. Why is this so?

Iron-Core Chokes

A choke is a coil used to purposely introduce a high reactance in a circuit. An iron-core choke is a coil wound on an iron core. A typical iron-core choke and the schematic symbol used to represent it are shown in Fig. 15.

Iron-core chokes are sometimes found in audio equipment, but their most important use is in power supplies. Here they are used in conjunction with capacitors to smooth the pulsating dc from the output of the rectifier. Chokes that are used for this purpose are called filter chokes or smoothing chokes.

PHYSICAL AND ELECTRICAL CHARACTERISTICS

It is sometimes difficult to distinguish an iron-core choke from a small power transformer. However, an iron-core choke usually has only two leads, whereas a power transformer has more than two leads. However, some iron-core chokes have a tapped winding. This type of choke has three leads and is very rare; most chokes have only two leads and can be distinguished easily from a transformer by this characteristic.

The core on which an iron-core choke is wound is similar to a transformer core. Thin sheets of laminated silicon steel are used in constructing the core. A frame is generally provided around the core and this frame serves the dual purpose of holding the laminations of the core together tightly and also in providing a convenient method of mounting the choke. You can see the frame in Fig. 15.

The core of an iron-core choke is made of silicon steel, which keeps the hysteresis losses in the choke as low as possible. The core is laminated to keep the eddy current losses low. The pulsating dc that is fed to a filter choke is actually a mixture of ac and dc. The ac has exactly the same effect in the choke as it has in a transformer, so both hysteresis and eddy current losses are present.

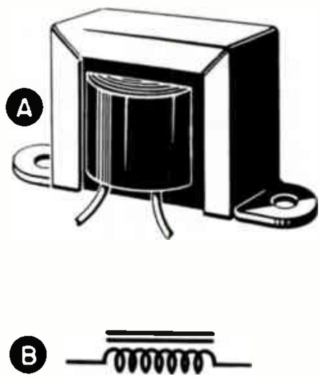


Fig. 15. A typical iron-core choke and its schematic symbol.

A choke used in a power supply usually has a fairly high inductance. Inductances from about 1 henry up to 30 or 40 henrys are quite common.

The size of wire chosen for winding a choke is determined by the current that will flow through the choke.

When an iron-core choke is used as the filter choke in the power supply, there is usually a high dc current flowing through it; superimposed on this dc is ac. The wire with which the choke coil is wound must be large enough to accommodate this current without overheating.

The current that flows through the choke coil will govern the size of core required. Saturation, which you have already studied, must be avoided in chokes. It can be avoided only by using an iron-core large enough to handle the magnetic field developed by the coil without reaching the saturation point.

If a choke becomes saturated, its inductance drops. This in turn results in a decrease in the inductive reactance and hence a decrease in its effectiveness in filtering the pulsating dc to pure dc.

In your career as an electronics technician, you will have occasion to replace defective filter chokes. When selecting a replacement, keep in mind that if you use too small a choke, it can be saturated and thus be very ineffective as a filter. Therefore, the replacement choke should be of approximately the same physical size as the original choke and the wire used to wind the choke should be at least as large as the wire used to wind the original choke. Technicians usually do not concern themselves with the wire size because manufacturers rate their chokes giving the inductance and the current that the choke is designed to handle. Thus, a choke rated at 8 henrys, 250-ma, is a choke that will have an inductance of 8 henrys when a dc current of 250 milliamperes is flowing through it. If the current flowing through the choke is higher than 250 milliamperes, the core will approach the saturation point, and the inductance will drop. If the current flowing through the choke is less than 250 milliamperes, the inductance of the choke will be somewhat higher than 8 henrys. This will not cause any trouble; the equipment will work as well as ever. If the current flowing through the choke exceeds the 250-ma current rating substantially, the chances are that the choke will overheat and may eventually get so hot that it will burn out.

HOW CHOKES ARE USED

We mentioned that chokes are used along with capacitors in power supplies to help smooth the pulsating dc at the output of the rectifier to pure dc at the output of the filter network. A typical circuit showing how a choke and a capacitor may be used is shown in Fig. 16. To see how the combination acts to filter the pulsating dc, first consider that the pulsating dc at the input actually consists of two components, ac superimposed on dc. To understand the action of the choke and capacitor, we can study their action on the two components separately.

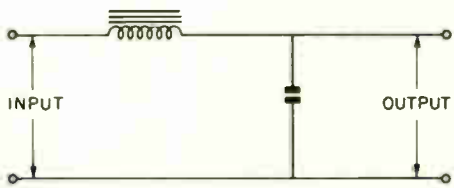


Fig. 16. A filter network made up of a choke and a capacitor.

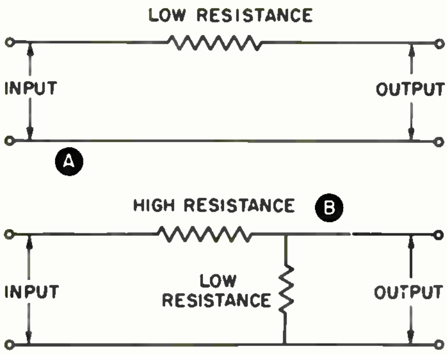


Fig. 17. The effect of the filter circuit for a dc component is shown at A, and for an ac component at B.

You know that a choke offers little or no opposition to the flow of dc through it. The only opposition the choke will offer to the flow of dc will be due to the resistance of the wire used to wind the coil. Since this is normally quite low, the dc can flow from the input through the choke coil to the output without any difficulty. The capacitor is charged by the dc, but once it is charged, it will not draw additional dc through the choke. Between the input and output circuits, we actually have a very simple circuit like the one shown in Fig. 17A insofar as the dc is concerned. Notice that we have only a low resistance in the circuit. The capacitor is not shown, and it can be completely ignored insofar as the dc flowing in the circuit is concerned.

The action of the choke and capacitor to the ac component is completely different. The choke, since it has inductance, has inductive reactance and is usually selected so that the inductive reactance will be quite high. The capacitor on the other hand is selected with a low capacitive reactance. Thus, insofar as the ac is concerned, we have a circuit like the one shown in Fig. 17B. Here we have a high resistance taking the place of the choke and a low resistance taking the place of the capacitor. Now these two "resistors" act like a voltage divider and most of the voltage will be dropped across the high resistance and very little will appear across the low resistance in the output. If the reactance of the choke is ten times the reactance of the capacitor, the ac component at the output would be approximately 1/10 the ac at the input.

If further filtering is needed to smooth the pulsating dc more than it can be smoothed by a single choke and capacitor, two chokes and two capacitors can be used as shown in Fig. 18. Here, if there is a 10-to-1 reduction of hum (which is what the ac is called since it produces hum in the output of the device) in each section, the total hum reduction in a two-stage filter network of the type shown in Fig. 18 would be 100. This means that if a 100-volt ac signal is applied to the input of the circuit, there will be a 1-volt ac signal at the output. On the other hand, dc applied to the input would flow through the filter network relatively unhampered.

In a two-stage filter network such as shown in Fig. 18, the first choke, marked L1, is called the input filter choke and the second one, marked

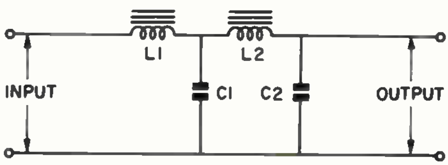


Fig. 18. A two-stage filter network.

L2, is the output filter choke. The choke, L1, is sometimes a special type of choke called a swinging choke. This type of choke is made with a rather small air gap in the core. The core is somewhat smaller than it should be for the amount of current that will flow through the choke so that it saturates rather easily and its inductance varies. The advantage of using this type of choke is that we are able to obtain better voltage regulation at the output of the power supply. By voltage regulation we mean keeping the output voltage more nearly constant as the load or current taken from the power supply vary. The second choke, marked L2, is usually called a "smoothing" choke.

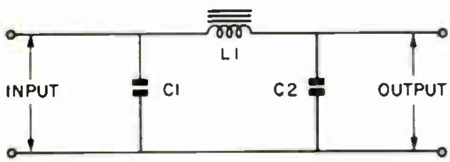


Fig. 19. A capacitor input filter.

Another type of filter network is shown in Fig. 19. This type of filter is often called a brute force filter. Notice that this network uses two capacitors and one choke. This type of network is called a capacitor input filter, whereas the one shown in Fig. 18 is called a choke input filter. In Fig. 19, C1 is called the input capacitor and C2 the output capacitor. The dc output voltage obtained from a capacitor input type of filter is somewhat higher than can be obtained from a choke input filter, but the voltage regulation is better with a choke input filter. Both types are found in modern electronic equipment.

SUMMARY

Chokes are important to the electronics technician, because he will encounter them in most pieces of electronic equipment that he is called upon to service.

Most chokes look like iron-core transformers except that they have only two leads, whereas a transformer has at least three and usually more leads. The cores of chokes are made of laminated sheets of silicon steel. This type of construction is used to keep the eddy current and hysteresis losses low.

Chokes can be saturated if the current passing through them is too high. Therefore, in replacing a defective choke in a piece of electronic equipment, the technician should use a replacement at least as large as the original. Filter chokes are used in conjunction with capacitors to smooth the pulsating dc at the output of a rectifier to pure dc at the output of the filter circuit. Chokes are used for this purpose because they offer a low-resistance path to the flow of dc through them but offer a high reactance to the flow of ac through them. The pulsating dc at the output of the rectifier actually consists of a dc component with an ac component super-imposed on it. The choke lets the dc component go through with little or no effect on it, and in conjunction with the capacitor greatly reduces the ac component.

SELF-TEST QUESTIONS

(ag) How can you tell a filter choke from a power transformer?
(ah) What is the purpose of the fil-

ter choke in a power supply?

(ai) What is the danger of passing too high a current through a filter choke?

(aj) Does a filter choke offer a high resistance or does it offer a low resistance to the flow of dc through it?

(ak) In a two-stage filter network, what are the names given to the two filter chokes?

(al) What is a swinging choke?

Relays

Another magnetic device is the relay. Although relays are used chiefly in transmitters and industrial equipment, they are found in some radio and TV receivers with automatic tuning systems. Relays are used to open and close circuits electrically. They are electric switches. In many ways a relay is similar to a mechanical switch. In order to understand relays, you should understand mechanical switches. So let's look at them first.

SWITCHES

Switches are made with several different contact arrangements. The simplest switch is one that has two positions. In one position, the circuit is open; in the other position the circuit is closed. This is called a single-pole, single-throw switch, abbreviated SPST. Fig. 20 shows an example of this kind of switch. In the position shown at A, the circuit is closed. When the blade is raised as at B, the circuit is open.

The switch illustrated in Fig. 20 is only one kind of SPST switch. They are made with different types of contact arrangements. However, although the mechanisms are different, the electrical principles are the same--in one position of the switch the circuit is open, and in the other position the circuit is closed. An ordinary light switch such as you have in your home is another example of an SPST switch.

Another type of switch, called a single-pole, double-throw switch (SPDT) is arranged so that in one

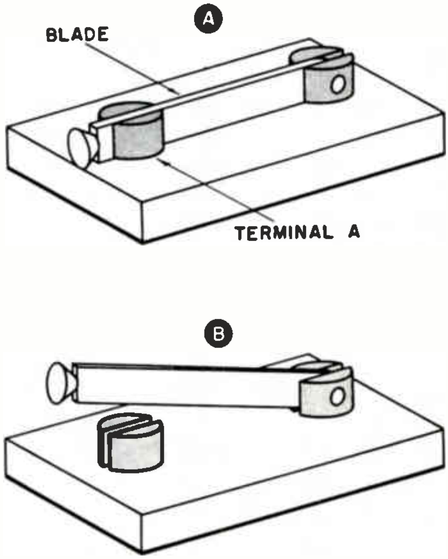


Fig. 20. A single-pole, single-throw switch is shown closed at A, and open at B.

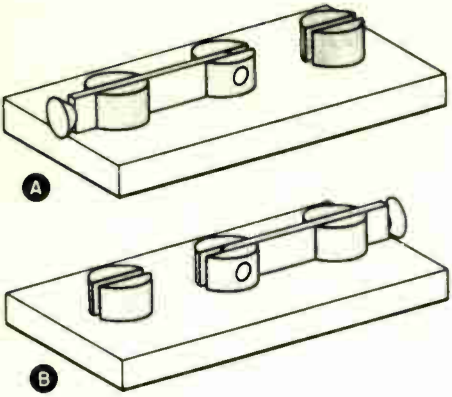


Fig. 21. A single-pole, double-throw switch makes contact in either of the two positions shown at A and B.

position it closes one circuit, and in the other position it closes another circuit. Fig. 21 shows an example of this type of switch.

A double-pole, single-throw (DPST) switch shown in Fig. 22 is really two switches in one. It has two blades, mechanically joined, so they are thrown open or closed at the same time. As in the SPST switch, when the blades are down, the circuits are closed, and when they are up, the circuits are open.

The double-pole, double-throw (DPDT) switch shown in Fig. 23 has

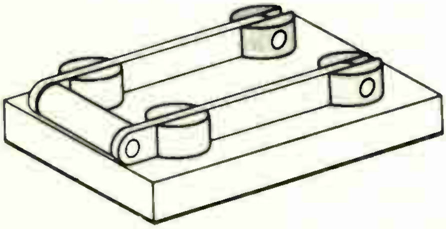


Fig. 22. A double-pole, single-throw switch.

two blades joined mechanically and two closed positions. There are also triple-pole, single-throw (TPST) switches, and triple-pole, double-throw (TPDT) switches.

Relays are also made with all these different types of contacts. Let's see how a relay works.

SIMPLE RELAY CONSTRUCTION

In its simplest form, a relay consists of nothing other than an iron-core coil such as shown in Fig. 24

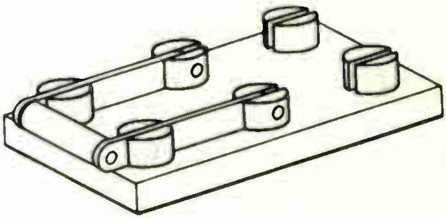
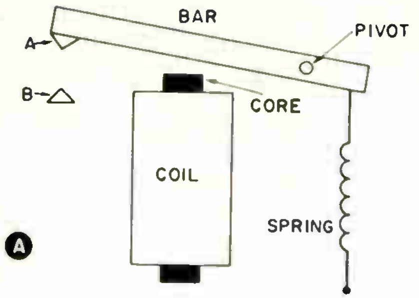


Fig. 23. A double-pole, double-throw switch.

with a bar of magnetic material placed on a pivot near one end of the core. One end of the bar is attached to a spring. The tension of the spring lifts the bar up and away from the core of the magnet. The motion of the top of the bar is usually restricted by some non-magnetic material so that when the relay is not energized, the bar will assume the position shown in Fig. 24A.

When a voltage is applied to the relay coil, current flows through the coil and the magnetic field produced attracts the bar on top of the coil and pulls it down as shown in Fig.



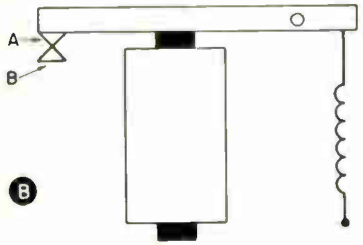


Fig. 24. Basic operation of a relay.

24B so that the contacts A and B are closed. Leads can be connected to terminals A and B so that energizing the relay closes the circuit and current will flow through the circuit.

Some relays operate from dc, but others are made that operate from ac. There is usually not too much difference between small ac and small dc relays, except that the spacing between the bar and the magnet is usually somewhat greater in an ac relay. If the bar comes too close to the magnet in an ac relay, there may be some tendency for the relay to chatter. By chatter, we mean that the bar moves up and down as the ac goes through its cycle and the strength of the magnet varies. If the bar is made of heavy enough material and kept a reasonable distance from the magnet, this problem is usually not encountered in small ac relays.

RELAY CONTACTS

The contacts on relays are identified using the same system that is used to identify the contacts on an ordinary switch. The movable arm of a relay or switch may be made to make contact in one position and no contact in the other position, or it may be made to make contact in either of two positions.

If there is only one set of contacts on a relay, we call it a single-pole, single-throw relay. The relay shown in Fig. 24 has this type of contact.

Single-pole, double-throw relays are arranged so that when the coil is energized, it pulls the bar down and the bar makes contact with one terminal and when the coil is not energized, then the bar moves up under the tension of the spring and makes contact with another terminal, as shown in Fig. 25.

Relays are also made with all the other contact arrangements that we

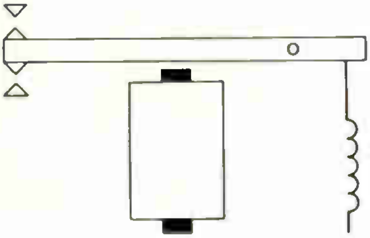


Fig. 25. A single-pole, double-throw relay.

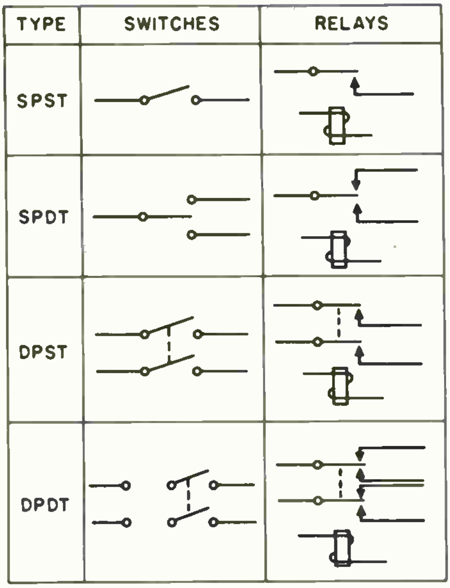


Fig. 26. Schematic symbols for different types of switches and relays.

showed for mechanical switches. Fig. 26 shows the schematic symbols for different types of switches and relays.

SPECIAL PURPOSE RELAYS

There are many special types of relays designed for specific jobs.

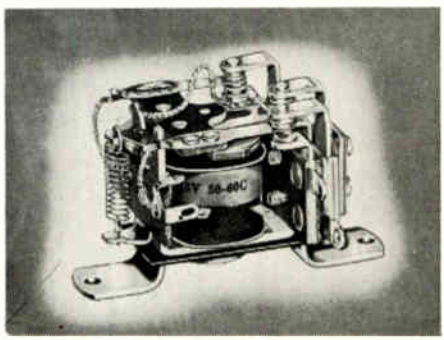
Time-Delay Relay.

One type of relay is a time-delay relay. In this type of relay, the contacts do not close until a predetermined time has elapsed after the power is applied to the relay. After this time has elapsed, the contacts are closed and the circuit is complete. Time-delay relays are usually either SPST or DPST relays. Double-throw time-delay relays are seldom used. Time-delay relays are often used in transmitting equipment where it is desirable to allow the cathode of the tubes to warm up before the other operating voltages are applied.

Overload Relay.

Another type of relay is used for protective purposes in electronic equipment. This type of relay is often called an overload relay. It is designed so that the relay is energized and the circuit opens when the current exceeds a predetermined value. Sometimes a break-down in one circuit will affect another. The current in the second circuit might rise so high that some valuable part such as a tube would be destroyed. With an overload relay in the circuit, when the current rises above a safe value, the relay automatically opens, removing the voltage from the tube, thus saving it from destruction.

As you might expect there are some circuits in which if the current drops below a predetermined value, some damage might result or the equipment might fail to operate



A relay of the type found in electronic equipment. This is a triple-pole, double-throw type.

Fig. 27. A thermal type overload device.

properly. There are relays that are set to open when the current drops below a predetermined value.

Another type of overload device that is widely used in TV receivers is shown in Fig. 27. This type of overload is not a relay, but instead is made of a metal strip which is made from two dissimilar metals. The current flowing through the strip causes the metals to heat and they expand at different rates. If the metal gets hot enough, the expansion will be great enough to distort the metal so that it springs loose from the contact holding it in place and this opens the circuit. You will run into this type of overload device frequently in TV receivers. Be sure not to confuse it with an overload relay. It serves the same basic purpose, but it is a thermal overload device rather than a magnetic relay overload device.

SUMMARY

Relays are devices used to close circuits automatically, to protect circuits either by waiting for a predetermined time before closing the circuit or by opening the circuit when an overload occurs. The relay consists of an electromagnet, a bar which can be held in one position by the magnet and returned to the other position by the spring, and one or more sets of contacts.

Relays are found in many different types of electronic equipment. About the only types of defects you are likely to encounter in relays are open coils, and burned or dirty contacts. Sometimes replacement coils are available, sometimes the contacts can be cleaned, but with some types of relays the only suitable remedy for a defect is to replace the entire relay.

SELF-TEST QUESTIONS

(am) What do we mean when we refer to a switch as an SPST switch?
(an) What is meant by SPDT?
(ao) What do the letters DPDT mean?
(ap) What causes the contacts in a relay to close?
(aq) What is a time-delay relay?
(ar) What is an overload relay?
(as) Overload protective devices are frequently found in TV receivers - are these overload devices overload relays?

LOOKING AHEAD

The iron-core devices that you studied in this lesson are all devices that will be found frequently in electronic equipment. Of the three, transformers and chokes will be found in more different types of equipment than relays, but there are many pieces of equipment that use a large number of relays. Defective transformers, chokes, and relay coils can frequently be recognized by a characteristic odor that they give off when they have been overheated. Also, you will generally notice signs of sealing wax and other compounds leaking from an overheated iron-core device. The insulation becomes brittle and can often be crumbled with your fingers. You will not go far in your electronics career before you run into one of these devices that has broken down. Usually it is a simple matter to locate and replace one of these parts that is defective.

In the lessons you have so far studied, you have covered a number of the basic components found in electronic equipment. Now, you are ready to study vacuum tubes, and then see how the components are used with vacuum tubes in amplifiers and other interesting circuits.

ANSWERS TO SELF-TEST QUESTIONS

(a) The gilbert is a unit of magnetomotive force. The gilbert is slightly smaller than the ampere-turn. The gilbert is particularly useful in expressing the magnetomotive force of a permanent magnet.

(b) The maxwell is the unit of magnetic flux. One line of flux is equal to one maxwell.

(c) The flux density is a measure of the number of flux lines passing through a given area. A flux density of one line or one maxwell per square centimeter is known as a gauss.

(d) 100 maxwells or lines. To find the total number of lines you multiply the area in centimeters by the flux density; in this case $10 \times 10 = 100$ lines.

(e) When magnetic saturation occurs, all of the molecules in the core material are aligned with their north poles pointing in one direction and their south poles pointing in the other so that any further increase in magnetomotive force cannot produce any further alignment of the molecules or any increase in flux.

(f) No - in fact it produces a number of undesirable effects and should be avoided.

(g) Magnetic saturation cannot be produced in air regardless of how strong the magnetomotive force is.

(h) A hysteresis loss is a loss due to the inertia of a magnetic circuit. When the core in a choke or transformer is magnetized by passing a current through the coil, the core does not return to a state of zero magnetism when the current is removed. Some magnetism will remain and power must be used to bring the flux density back to zero. This power required to bring the flux density back to zero represents the hysteresis loss.

(i) Eddy current losses are losses due to the fact that a core of the transformer choke acts like a complete turn of a coil and the flux cutting this turn induces a voltage in it which causes a current to flow. This represents a loss which is known as an eddy current loss.

(j) By selecting a material that has very little magnetic inertia. Materials such as silicon steel have much lower hysteresis loss than hard steel.

(k) Eddy current losses are kept at a minimum by making the core of a transformer or choke coil in the form of sheets of metal rather than a solid piece. The sheets are electrically insulated from each other so that the flow of eddy currents across the entire core material is prevented.

(l) Flux leakage losses are due to the fact that part of the flux escapes from the core and travels through the air surrounding the core. This flux serves no useful purpose since the flux lines in the case of a transformer would not cut the secondary winding of the transformer.

(m) In its simplest form a transformer is two separate coils of wire wound on a common core and placed near each other so that the magnetic lines produced by one coil will cut the other coil.

(n) The two types of losses encountered in a transformer are core losses and copper losses.

(o) The power lost in a transformer is turned into heat. The heat is radiated by the transformer and the chassis on which the transformer is mounted.

(p) Yes. A transformer is one of the most efficient devices you will encounter in electronics. Large power transformers may have efficiencies as high as 98% or better. This means that 98% of the power taken by the primary winding of the transformer is available for useful work at the secondary winding of the transformer.

(q) No. A transformer designed for 60-cycle operation will overheat and burn out if it is operated on a 25-cycle power line.

(r) A step-up transformer is a transformer that has more turns on the secondary winding than on the primary winding. As a result, the voltage available across the output of the secondary winding will be higher than the voltage supplied to the primary winding. The ratio by which the voltage is stepped up is determined by the ratio of the number of turns on each winding of the transformer. In other words, if the secondary winding has twice as many turns as the primary winding, the voltage available across the secondary will be twice the voltage applied to the primary.

(s) 1 amp. A step-down transformer steps the voltage down by the turns ratio; however, it steps the current up by the same turns ratio. Therefore, if the secondary current is 3 amps, the primary current will be only 1 amp.

(t) 2 amps. To find the current we divide the power by the voltage; in this case we have 230 watts divided by 115 volts equals 2 amps. The efficiency of a transformer is usually so high that we can ignore any losses and if the power consumed by the secondary is 230 watts, we can consider that the primary power will also be 230 watts.

(u) If the turns ratio of a transformer is 2:1, and the transformer is operated on a 120-volt power line, then the secondary voltage must be 60 volts. Since the current drawn by the secondary winding is 4 amps, then the power the secondary winding is supplying is $60 \times 4 = 240$ watts. The primary winding must take this much power from the power line and since the voltage at the power line is 120 volts then the current must be 240 watts divided by 120 volts = 2 amps.

(v) No. Step-up type audio transformers were used in the early days of radio when the gain that could be obtained with a single vacuum tube was comparatively low. By using a step-up type transformer some increase in voltage could be obtained in the transformer and this helped obtain a reasonable voltage amplification in the stage. However, with modern tubes and transistors, adequate gain can be obtained in a stage without resorting to step-up transformers.

(w) Audio transformers are most widely used in electronic equipment as impedance-matching devices. They are used to match the impedance of one stage to the impedance of the following stage, and they are also used to match the impedance of the output stage to the loudspeaker voice coil impedance.

(x) The output stage of a radio or TV receiver must be matched to the speaker voice coil in order to get maximum power transfer from the output stage to the speaker voice coil.

(y) A transformer with a tapped secondary. The center tap on the secondary is either grounded or used to feed bias to the push-pull stages and the signals supplied to the stages are 180° out of phase, so when the current is increasing in one stage of a push-pull amplifier it is decreasing in the other.

(z) The turns ratio must be 10:1. The turns ratio is equal to the square root of

$$\frac{Z1}{Z2}$$

Substituting 1000 ohms for Z1 and 10 ohms for Z2 we find that the turns ratio is

$$\frac{1000}{10} = \sqrt{100}$$

$$\sqrt{100} = 10$$

(aa) An autotransformer is a transformer that consists of a single winding with a tap.

(ab) An autotransformer can be either a step-up transformer or a step-down transformer. If the primary voltage is applied across the entire winding and the secondary voltage taken off between the tap and one of the primary connections, the transformer will be a step-down transformer. On the other hand, if the primary voltage is applied between the tap and one of the outside connections and the secondary voltage taken off across the entire winding, the output transformer will be a step-up transformer.

(ac) The disadvantage of an autotransformer is that the primary and secondary windings are not completely isolated from each other electrically.

(ad) An rf transformer is a radio frequency transformer used between radio frequency amplifier stages.

(ae) RF transformers have either an air-core or a powdered iron core.

(af) The secondary winding of an i-f transformer, along with the capacitor connected across it, form a series-resonant circuit. This results in a high circulating current and a resonant voltage step-up across the coil and across the capacitor. As a result, the voltage across the coil will be considerably higher than the source voltage. We refer to this increase in voltage as resonant voltage step-up and it explains why the voltage across the secondary winding of an i-f transformer is higher than the voltage across the primary winding.

(ag) A filter choke has only one winding and therefore it has only two leads. Transformers usually have two or three windings and therefore will have more than two leads.

(ah) The filter choke is used in conjunction with filter capacitors to help smooth the pulsating dc from the rectifier output to pure dc.

(ai) The excessive current may over-heat and burn out the wire used to wind the filter choke. In addition, the filter choke may become saturated, even if it doesn't burn out, due to the excessive current flowing through it and its effectiveness as a filter will be reduced.

(aj) A filter choke offers a low resistance to the flow of dc through it. The only opposition the choke offers to dc is the resistance of the choke which is equal to the resistance of the wire used to wind the choke. On the other hand, a filter choke offers a high opposition to the flow of ac through it.

(ak) The two filter chokes are known as the input choke and the output filter choke. The choke connected to the rectifier is the input filter choke, and the other choke is the output filter choke.

(al) A swinging choke is a choke made with a small air gap in the core. The choke is usually somewhat smaller than it would normally be for the amount of current flowing through it and as a result the choke saturates easily. Its inductance varies as the current flowing through it changes.

(am) SPST means single-pole single-throw. With this type of switch there is only one circuit and in one position the switch is closed and in the other position it is open.

(an) SPDT means single-pole double-throw. With this type of switch one circuit can be completed with the switch in one position and a second circuit can be completed with the switch in the other position.

(ao) DPDT means double-pole double- throw.

(ap) A current flowing through the relay coil produces a magnetic field which attracts a bar which is pivoted above the magnet. The magnet pulls the bar towards the magnet and the the bar closes the relay contacts.

(aq) A time-delay relay is a relay in which the contacts do not close when the relay is at first energized until after a predetermined time has elapsed. This type of relay is frequently found in transmitting equipment - its purpose is to permit the tubes to come up to operating temperature before the high voltages are applied to the tubes.

(ar) An overload relay is a relay used for protective purposes. The relay contacts are normally closed, but if the current flowing through the relay coil exceeds a certain predetermined value the relay contacts will open protecting the circuit in which the relay is used.

(as) No; overloads of this type are thermal overloads and operate when the current flowing through them causes the metal conductor to reach a certain temperature. This causes the metal to spring the contacts open. An overload device of this type is called a thermal overload and is not a relay.

Lesson Questions

Be sure to number your Answer Sheet B109.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What do we mean when we say a magnetic circuit is saturated?

2. Name the three types of core losses encountered in iron-core transformers.

3. If the primary of a 1:3 step-up transformer is connected to a 120-volt power line, what voltage will be produced across the secondary?

4. If we connect a 200-ohm resistor across the secondary of a 1:2 step-up transformer, and then connect the primary to a 100-volt source, what will the primary current be?

5. What is the big disadvantage of interstage audio transformers?

6. What must the turns ratio be of the matching transformer used to match a tube with an output impedance of 1600 ohms to a 16-ohm speaker?

7. Why are transformers with separate primary and secondary windings more suitable than auto transformers in many applications?

8. Could a 10-henry, 250-ma choke be used as a replacement for a 10-henry, 200-ma choke?

9. Why is a swinging choke often used as the input choke in a two-stage filter network?

10. What does the abbreviation DPDT mean?

DETERMINATION

Did you ever watch a modern, diesel earthmover-- one with 10 times the horsepower of a car and rubber tires as tall as a man? Wasn't it a thrill to see that great machine level obstacles in its path, leaving a perfectly smooth roadway? The word "determination" always brings to mind that picture of an earthmover -- a machine which goes places once the throttle lever is thrown.

So, too, are you determined to go places, to achieve success and happiness. One by one you are completing your lessons, studying hard and making sure you understand everything; step by step you are approaching that greatest of all goals -- SUCCESS.

Of course, the way is long and not always easy. But whenever the going gets a little tougher than usual or you feel a bit discouraged, just bring out that old determination, and back it up with every single ounce of ambition you have.

J E Thompson

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-209

ACHIEVEMENT THROUGH ELECTRONICS

nri

HOW VACUUM TUBES WORK

B110

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

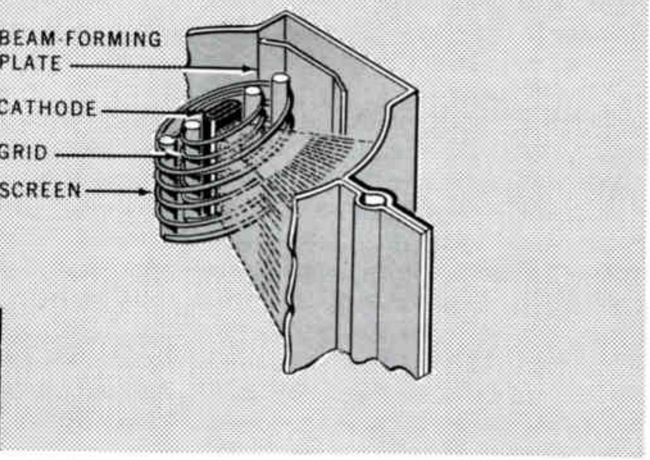
HOW VACUUM TUBES WORK

B110

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

JDL3M1174 1974 EDITION Litho in U.S.A.



HOW VACUUM TUBES WORK

Vacuum tubes are used not only in radio and TV entertainment broadcasting, but also in commercial radio equipment, telephone repeater systems, diathermy equipment, and in many types of industrial electronic equipment. Although transistors have taken over many of the jobs originally done by tubes, there are still many applications where the vacuum tube is superior to the transistor. Even in those applications where the transistor is superior to the vacuum tube, there are many older pieces of electronic equipment still in use using vacuum tubes which you will have to service. There are other applications where a transistor could be used in place of a vacuum tube, but vacuum tubes are used because they cost less than transistors or they can be manufactured in large quantities to closer specifications than transistors can be manufactured.

There are many different sizes and shapes of vacuum tubes. Some tubes are so small they are scarcely larger than a thumbnail; other tubes found in large radio and TV transmitters are several feet tall. Regardless of whether the tube is a miniature tube such as those found in hearing aids or a large tube such as those found in transmitters, it works on the same basic principles. In this lesson you will study tubes in detail. Later you will see how these tubes are used along with the parts you studied in previous lessons.

ELECTRON EMISSION

In the circuits you studied in previous lessons, the electrons stayed within the circuit wiring and flowed only over a solid path. For example, when you connect a resistor across a terminal of a battery, the electrons are set in motion around the circuit. Electrons flow through the resistor, the wires and the battery, but they stay within the closed circuit. Electrons do not leave the wire connecting the resistor to the battery and travel off into space around the wire, nor do they leave the resistor and move into the space around it. In a tube, however, electrons are forced to fly off into the space surrounding one of the elements in the tube. This element is called the cathode. When an electron leaves the cathode we say it is "emitted" by the cathode; the giving up of electrons by the cathode is called "emission".

You will remember that all material is made up of atoms and that one of the parts of the atom is the electron. The electrons in an atom are in a continuous state of motion. The speed and the amount of motion depend, for one thing, upon the temperature of the material. Normally the atomic force within the atom prevents electrons from escaping and flying off into space. This is true even in the case of the free electrons in a conductor that move through the conductor when current flows. However, if enough heat is supplied to a conductor, it is sometimes possible to overcome the force holding the electrons within the surface of the conductor and to drive some of the electrons off into the space surrounding it. This is what we mean by emission. This type of emission is often called "thermionic" emission; thermionic means emission by heat. Some materials give off electrons more readily and at lower temperatures than other materials. Often the cathodes of tubes are coated with these materials that will readily give off electrons at low temperatures.

Altogether there are four ways in which electrons can gain enough energy to escape from a material into the space around it. These are: (1) they can be evaporated or driven out by applying heat; (2) they can be driven off by bombardment by very small, high-speed particles such as other electrons; (3) they can be driven out of some materials by the energy in light rays; and (4) they can be jerked or pulled out by a very high positive potential placed on a nearby metallic object. All four of these methods are used in various types of electron tubes to provide the free electrons that all tubes depend upon for their operation. However, the first method, evaporation by heat, is by far the most important, and we will spend more time in this lesson on tubes using this type of emission. We will also briefly discuss the second and third types of emission.

Thermionic Emission.

As we mentioned, when electrons are driven from a metal or metallic compound by means of heat, this type of emission is called thermionic emission. Thermionic is pronounced THERM-I-ON-IK. As we mentioned, thermo means heat, ionic refers to electrons and hence the word thermionic is used to describe the type of emission where the electrons are driven from the cathode by heat.

In the operation of a vacuum tube it makes no difference where the heat comes from; if the cathode can be made hot enough it will emit electrons. However, the most convenient method of producing the heat needed is by means of a heater or filament placed inside the vacuum tube. A voltage is applied to the heater or filament and this voltage causes a current to flow through the heater or filament and causes it to heat to at least a red heat, and sometimes to a white heat. The hot filament or cathode then gives off the electrons required to operate the tube. It is important to realize the heater or filament voltage applied to the tube for the purpose of heating the tube does not actually enter into the operation of the tube in any way except to provide the energy necessary to heat the cathode. If instead of using electrical energy to heat the cathode we were able to heat the cathode with a gas flame, the heater or filament voltage could be removed entirely and the tube would perform just as well.

The terms heater, filament and cathode might be somewhat confusing at this time. A cathode is an electrode which gives off electrons. It is heated by means of an element called a heater. A filament is a combination heater and cathode. We will go into this in detail in the next section of this lesson.

Secondary Emission.

When electrons are driven off a metal by bombarding the metal with high-speed particles such as other electrons, we refer to this type of emission as secondary emission. What happens in this type of emission is that a particle travelling at a very high speed strikes the metal object with such force that it is able to dislodge a number of electrons from the material. These dislodged electrons fly off the material into the space surrounding it. You will see later that this type of emission is not always desirable; as a matter of fact, it creates a problem in some vacuum tubes.

Photoelectric Emission.

When electrons are driven out of the material by the energy in light rays, this type of emission is called photoelectric emission. Photoelectric tubes are used in the motion-picture industry in connection with the sound track. A sound track is put on the side of the film. Light passing through this sound track strikes a photoelectric tube. The density of the sound track varies as the speech or sound originally recorded on the film varies. This causes the amount of light striking the photoelectric tube to vary, which in turn varies the number of electrons emitted.

Photo-emission is not often encountered in radio and TV servicing, but it is of importance to the industrial electronics technician. Even the radio-TV serviceman should know something about it, because photoelectric tubes have been used in the past in radio-phono combinations. In addition, you may have occasion to service the sound system of a home movie projector.

Now let's see how tubes work. You already know something about how a vacuum tube operates. However, in the explanations we have given you previously, we have left out many details in order to approach the subject gradually. Now we will learn more about tubes by studying all these details. Let us start by studying the simplest tube.

The Diode Tube

The simplest tube has only two elements, one to give off the electrons and another to receive them. The element that gives off the electrons is called the cathode, and the element that receives the electrons is called the plate or anode. A simple tube having only two elements is called a diode.

Other tubes have other elements in addition to these two. Therefore as we study the diode, remember that tubes with more than two elements are really diodes with additional elements added.

TUBE CATHODES

Tube cathodes can be divided into two types, those that are directly heated and those that are indirectly heated. Directly heated cathodes are called filament-type cathodes or, more frequently, simply filaments. This type of cathode was used in early vacuum tubes and is still used in tubes designed for battery operation, and in large transmitting tubes. Indirectly heated cathodes are simply called cathodes.

Filament-Type Cathodes.

The schematic symbol used to represent a filament type of cathode is shown in Fig. 1. The voltage used to heat the filament is applied directly between the two leads from the filament.

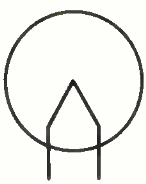


Fig. 1. Schematic symbol used to represent a filament type of cathode.

In large transmitting tubes the filament is made either of pure tungsten or of a mixture of tungsten and thorium. The very large transmitting tubes generally have a pure tungsten filament. Tungsten is a metal that can be operated at very high temperatures. Most electric light bulbs manufactured today have a large percentage of tungsten in the filament that is heated to a very high temperature to give off light.

The filaments used in many of the smaller transmitting tubes are made of a mixture of thorium and tungsten and are called thoriated filaments. The addition of thorium to the tungsten provides a material that will give off electrons at a somewhat lower temperature than pure tungsten. Thus the amount of power required to heat the filament is lower than for a pure tungsten filament. Thoriated tungsten is not as suitable as pure tungsten in large transmitting tubes. The very high voltages used on these tubes can pull the thorium right out of the filament and thus destroy the tube. In smaller transmitting tubes the voltages used are not high enough to do this.

Filament-type receiving tubes designed for operation in portable receivers were widely manufactured at one time. The filaments of these tubes were coated with oxides of certain metals. This type of filament is called an oxide-coated filament and it has the characteristics of giving off electrons at a still lower temperature than a thoriated filament. Thus the filament power required by this type of tube is even less than that required by the thoriated filament.

Even though filament-type receiving tubes operated on comparatively low voltages and required only a very small current, the fact that the heater power served no useful purpose led to the disappearance of this type of tube. Modern portable receivers all use transistors since no filament power is required. However, you may occasionally run across an older portable receiver that some set owner is particularly fond of and have to fix this receiver. Generally, filament-type tubes are available for replacement purposes.

Oxide filaments are found in some small transmitting tubes used in mobile applications. However, if the voltage applied to the tube exceeds 500 volts by very much, the oxide on the filament may be pulled off by this voltage and therefore oxide-coated filaments will be used only in small transmitting tubes.

The filaments of transmitting tubes that are made of either tungsten or thoriated tungsten can be operated on either ac or dc. However, the oxide-coated filament used in small tubes designed for use in battery-operated equipment are usually made very small and thin in order to keep the filament power required as low as possible. If these filaments are operated from ac, as the ac drops to zero and then rises to a maximum value twice during each cycle, the current flowing through the filament will vary, causing the temperature of the filament to vary. This will cause a variation in emission from the filament resulting in hum. Therefore the filament of small transmitting tubes and the older obsolete portable receiver types must be operated from dc.

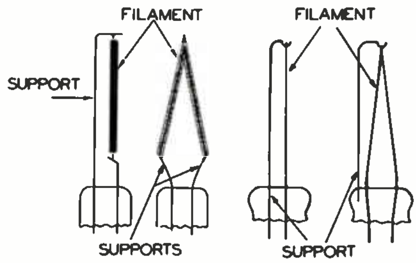


Fig. 2. Typical filaments.

The filament of a vacuum tube must be supported so it will stay in position. Typical filaments showing the type of support used are shown in Fig. 2. It is important that the filament be held tight so that it cannot sag and short to nearby elements in the tube. As a matter of fact, if the position of the filament changes, even though the filament may not touch any other elements in the tube, the characteristics of the tube will change because there are several tube characteristics that depend upon the spacing between the filament and the other elements in the tube.

Indirectly-Heated Cathodes.

The schematic symbol used to represent an indirectly heated cathode is shown in Fig. 3. The indirectly heated cathode is simply called a cathode. Notice that in addition to the cathode we have another element drawn beneath the cathode in the schematic symbol. This is the

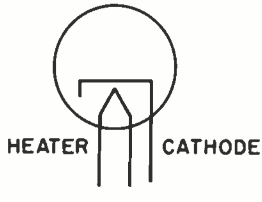


Fig. 3. Schematic symbol used for a heater and indirectly heated cathode.

heater, which is used to heat the cathode. Sometimes this is loosely called a filament because of its similarity to the filament we have just discussed and because the first tubes had directly heated cathodes. In fact the transformer winding used to supply heater voltage is still often called the filament winding.

The cathode is built in the form of a hollow cylinder like those shown in Fig. 4. The heater is placed inside of this hollow cylinder. Voltage is applied to the heater, and the heat produced by the heater is radiated and in turn heats the cathode.

The cathode is usually coated with oxide. This is done in order to provide an abundant supply of electrons at low operating temperatures. This type of cathode is used only in receiving tubes or small transmitting tubes where the applied voltage is not high enough to pull the oxide material off the cathode.

In many schematic diagrams of circuits in which an indirectly heated cathode type of tube is used, the heater is omitted. The heater actually serves no useful purpose as far as the operation of the tube is concerned, other than to heat the cathode. It does not enter into the characteristics of the tube and therefore the heater connections can be omitted to simplify the schematic diagram. The connections to the other elements in the tube are the ones that actually determine what the tube will do and how it will operate. We will follow this practice in many cases so we can emphasize the operating circuits.

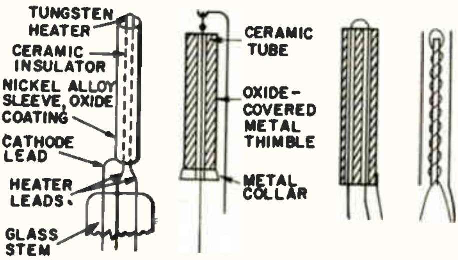


Fig. 4. Indirectly heated cathodes.

Operating Voltages.

The filament or the heater of a vacuum tube is designed to operate on a certain definite voltage. The first number used to identify receiving tubes gives an indication of the heater or filament voltage. For example, a 12L6 tube operates on a heater voltage of approximately 12 volts. The exact voltage is 12.6 volts. The number preceding the first letter indicates the heater voltage. A 35Z5 tube requires a heater voltage of 35 volts. A 6F6 tube requires a heater voltage of 6.3 volts, and a 2BN4 tube requires a heater voltage of 2.3 volts. Keep this in mind; it will be helpful to you when you start doing service work. The first number or group of numbers preceding the letters in the tube designation is an indication of the filament or heater voltage for which the tube was designed. This system is followed only for modern receiving tubes. Older receiving tubes did not use this system, and transmitting and industrial tubes do not use it.

THE PLATE

In the first vacuum tubes made, the anode that received electrons emitted by the cathode was simply a flat piece of metal, and hence was called a plate.

In modern vacuum tubes the plate completely surrounds the filament or cathode. The shape of the plate

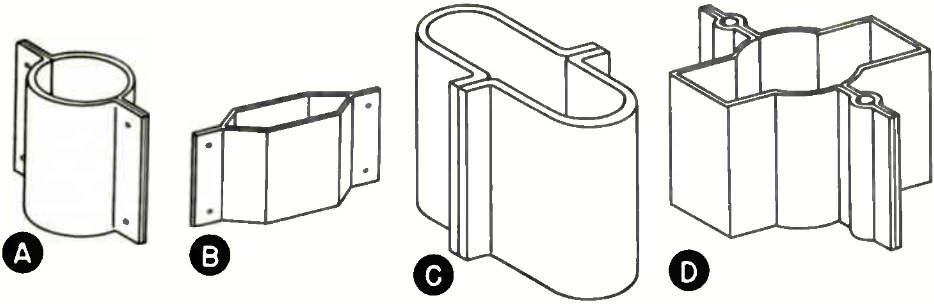


Fig. 5. Typical plate structures.

depends to some extent on the type of cathode used. If the cathode is simply a round cylinder, then the chances are that the plate will also be a round cylinder. However, if a filament type of cathode is used, such as those shown in Fig. 2, the plate will usually be rectangular in shape. Typical plate structures are shown in Fig. 5.

Vacuum tube plates may be made of any of several different materials, such as nickel, molybdenum, carbon, iron, tungsten, tantalum, and graphite. Zirconium is also sometimes applied as a coating on the plate. The plates of most receiving tubes and small transmitting tubes are made of nickel, which can easily be formed into the desired shape.

The plate of a vacuum tube is subjected to a certain amount of heating. Part of this heat comes from the filament or cathode, and part of it is produced by the electrons striking the plate. These electrons striking the plate can cause considerable heating; as a matter of fact in some transmitting tubes they produce so much heat that the plate becomes a bright red color.

The fact that the plate is heated results in a few additional problems. First, if the plate gets hot enough it will give off electrons, and if this happens the tube may not work properly because electrons may be able to travel both ways through it. The tube would simply act like a resistor. Therefore steps must be taken to keep the plate below the temperature where it will emit electrons. This is not too big a problem in small receiving tubes, but in the larger tubes it can be quite serious. The plate is often given a dull black finish because a black surface radiates heat readily and therefore will be cooler than a polished surface. The plates of many transmitting tubes are fitted with fins to provide a larger surface to dissipate or get rid of the heat produced at the plate of the tube. Large transmitting tubes are often water-cooled. As a matter of fact, in some large broadcast stations the water used to cool the tubes in the transmitter is used to heat the building in the winter.

Even if the plate is kept cool enough to prevent thermionic emission, the electrons travelling from the cathode of the tube over to the plate can pick up enough speed to strike the plate so hard that they will knock other electrons loose from the surface of the plate. Remember, this is one of the types of emission we discussed before and it is called secondary emission. The electrons that are knocked off the plate of the tube in this way leave the plate at a rather low speed. The plate normally has a positive voltage applied to it and the electron is negative, so in a diode tube, if the plate is positive, the electrons are simply attracted back to the plate. Secondary emission does not cause any difficulty in a diode. However, you will see later that in tubes having additional elements, secondary emission can be a problem.

Another problem that is created when the plate of a tube becomes very hot is that almost all metals have a certain amount of gas trapped right in the metal. When a metal plate becomes very hot, the gas trapped in it may be forced out of the metal and into the space surrounding the plate of the tube. Gas in a tube can destroy its usefulness. Now let's see why it is important that the amount of gas in a tube be kept as low as possible.

GAS EVACUATION

Tubes are called vacuum tubes because all the gases inside the tube including air are normally evacuated. There are two important reasons why these gases must be removed from inside the tube. First, if air is permitted in the tube, the filament or heater will oxidize; in other words, it will simply burn up when heated. Secondly, even if this problem could be overcome, there is another important reason why all gases must be removed from the tube. Gases are made up of molecules. Although these particles are extremely small, they are nevertheless many times the size of an electron. An electron traveling from the cathode of the tube to the plate moves at a fairly high speed. The chances are that the majority of electrons traveling from the cathode to the plate would strike one or more gas molecules if there were a large amount of gas in the tube. If a high speed electron strikes a gas molecule the electron will be deflected from its path and it will knock electrons out of the gas molecule.

Normally a gas molecule has no charge. However, if it is struck by an electron, which knocks other electrons out of the molecule, the molecule will be short of electrons, and hence will have a positive charge. We call positively charged molecules "ions". These ions are large and heavy in comparison to electrons and have a fairly high positive charge on them. Since the cathode of the tube is normally connected to a negative voltage source, it is negative, and will attract these positive ions. As a matter of fact, the positive ions will pick up a fair amount of speed traveling to the cathode and may bombard it with such force that small particles of the cathode material will be knocked loose.

Gas inside a vacuum tube is extremely undesirable. Therefore, in the manufacturing process, every effort is made to remove all the gases from inside the tube. However, some of the gas will remain in the tube and additional gas will boil out of the materials from which the tube is made the first time the tube is heated. To get rid of these gases left in the tubes a "getter" is placed inside the tube.

A getter is a small cup containing chemicals. During the manufacturing process, the tube is first evacuated by means of pumps. This will remove most of the gases from the tube. The tube is sealed and then it is heated. In the heating process gases are driven off the metals inside the tube such as the cathode and the plate. At the same time, the getter is heated and the chemicals in the getter combine with any gas molecules released, forming metal compounds, which are deposited on the glass envelope of the tube. The compounds in the getter hold onto the gas molecules and will not easily release them into the space inside the tube again. The silvery appearance that many tubes have near the base of the bulb is produced by the compounds forming on the glass envelope of the tube.

By using this procedure it is possible to obtain an excellent vacuum in tubes with a pure tungsten filament. The vacuum inside a tube with a thoriated tungsten filament is not quite as good, and the vacuum inside tubes with oxide-coated filaments is still poorer. The reason for this is that you can heat a tube with a pure tungsten filament to a higher temperature than you can the other types and hence the action of the getter is even more complete in these tubes than it is in the other types.

For this reason, tubes with thoriated tungsten filaments or oxide-coated cathodes are limited to uses where the operating voltages are somewhat lower than those that can be applied to pure tungsten filament tubes.

The leads connected to the various elements inside the tube are brought through glass seals. The glass is heated to a high temperature and it flows around the leads, providing a nearly perfect seal. Thus once the tube is evacuated, it is almost impossible for air to leak back into the tube. We say almost impossible because there is no such thing as a perfect seal and air will gradually leak back into the tube. It may take several years for enough air to get into the tube to affect its performance, but if the tube is left around unused long enough, eventually enough gas will get into the tube so that its operation will be impaired.

Tubes that have had all the gas removed from inside them are called hard tubes. Most tubes found in radio and TV receivers are hard tubes. However, there is another group of tubes into which certain types of gas have been deliberately introduced. Mercury is put inside some diode tubes. When these tubes are operated, the mercury will heat and vaporize, filling the inside of the tube with mercury vapor. This type of tube is called a soft tube or a gaseous tube. Other gases are sometimes used, but mercury vapor is the gas you will be most likely to encounter. It is easy to identify a mercury vapor tube because mercury gives off a characteristic blue glow. Therefore, if you see a tube operating with a bright blue glow, the chances are it is a mercury vapor tube.

If there is excessive gas inside a hard tube it also will have a blue glow. However, the blue glow is not as bright as it is in a mercury vapor tube. Mercury vapor tubes are almost always diode tubes, although a special type called a thyratron has three elements and hence is a triode. These tubes are usually quite easy to pick out. If you discover a blue glow between the elements of a hard tube, the tube is gassy and should be replaced.

Before leaving this subject there is one other point that should be brought out. Electrons emitted from the cathode of a hard tube frequently travel at a high speed and miss the plate of the tube and strike the glass envelope. When they do this there will often be a blue glow on the envelope of the tube. This does not indicate a defect in the tube. If a hard tube shows this blue glow on the envelope of the tube you can forget about it; it does not mean anything. However, if the blue glow appears between the elements of a hard tube, the tube is gassy.

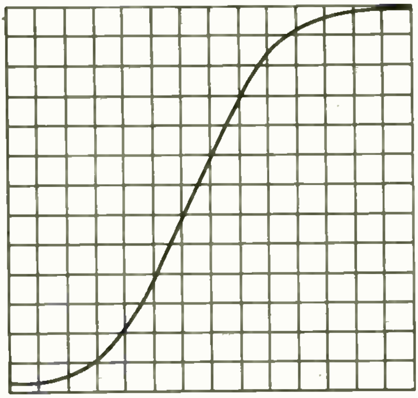


Fig. 6. Characteristic curve of a diode showing the relationship between plate current and plate voltage. Notice that there is a small plate current even when the plate voltage is zero. This is caused by a few electrons being emitted by the cathode at such a high speed that they travel over and strike the plate even when no voltage is applied to it.

CHARACTERISTIC CURVES

Among the information published by vacuum tube manufacturers are the characteristic curves of their tubes. These curves make it possible for the engineer or technician to predict how the tube will perform under a given set of operating conditions. The characteristic curve of a diode tube shows how much current will flow when a given voltage is applied to the plate of the tube. A typical diode characteristic curve is shown in Fig. 6.

Notice that this characteristic curve is not a straight line. It is bent on the two ends. Also notice that there is a very small current flow even when the plate voltage is zero. This is because a few of the electrons emitted by the cathode travel with sufficient velocity to reach the plate even without a positive voltage on the plate. As the plate voltage is slowly increased from zero to a high positive value the number of electrons flowing to the plate gradually increases. At first the increase is non-linear, but as the plate voltage is increased still further, eventually a point is reached where the characteristic curve becomes quite linear. By this we mean the curve is a straight line which indicates that we will get an almost constant change in plate current for a given change in plate voltage. For example, on a linear curve, if increasing the plate voltage from 25 to 50 volts causes an increase in current of 5 milliamperes, we can expect an increase in plate current of 5 more milliamperes when we increase the plate voltage from 50 volts to 75 volts. Again, if we increase the voltage from 75 volts to 100 volts and the curve is linear, we can expect another 5 milliampere plate current increase.

Notice that the top of the curve begins to round out and become flat--this is called plate current saturation. Eventually a point is reached where all of the electrons emitted by the cathode are drawn immediately to the plate of the tube. In other words, the electrons do not form a space charge or electron cloud around the cathode, but instead travel immediately from the cathode of the tube right over to the plate. When this point is reached, increasing the plate voltage still further will result in no further increase in the plate current flowing in the tube because the plate is pulling all the free electrons over to it and is gathering them up as fast as the cathode can emit them. As a matter of fact, if the plate voltage is increased beyond this point, there is the danger that some electrons will be emitted from the cathode by the process of jerking them out of the cathode by the high plate voltage. This may cause small particles of the cathode material to jerk loose, and if this happens the cathode of the tube will soon disintegrate and the tube will no longer be usable.

SUMMARY

There are a number of important things you should remember in this section. First, the cathode found in tubes can be divided into two types--the directly heated cathode which is called a filament, and the indirectly heated cathode, which is simply called a cathode. Remember that the heater used to heat an indirectly heated cathode performs no useful purpose other than to heat the cathode of the tube. It does not enter into the electrical circuit and operation of the tube itself.

Three types of cathode material are found in vacuum tubes: pure tungsten, thoriated tungsten, and oxide coatings.

Several different types of plates are used in vacuum tubes; the exact shape of the plate is usually determined by the shape and type of cathode used in the tube. The plate of a vacuum tube will give off electrons by thermionic emission if it becomes too hot. However, some transmitting tube plates can operate at a red temperature without giving off electrons, the materials used in the manufacture of the plates of these tubes are selected because they give off electrons only at a very high temperature. The plate may also give off electrons due to secondary emission. In some tube types this can become a problem.

The inside of a vacuum tube is normally highly evacuated. A tube with a high vacuum is called a hard tube. Gas is deliberately introduced into some tubes, and these are called soft tubes.

SELF-TEST QUESTIONS

(a) Name the two elements found in a diode tube.
(b) Into what two types can tube cathodes be divided?
(c) What is the name given to a directly heated cathode?
(d) How is the cathode of an indirectly heated tube heated?
(e) Why is the cathode of a receiving-type tube usually coated with an oxide?
(f) What useful purpose does the heater of an indirectly heated tube serve other than to heat the cathode of the tube?
(g) Approximately what would you expect the heater voltage to be of a type 8AC9 tube?
(h) What are the two names given to the element in the diode tube that receives the electrons?
(i) What two things cause the plate of a vacuum tube to be hot?

(j) What two undesirable things may happen if the plate on a vacuum tube becomes excessively hot?
(k) What is the purpose of a getter inside a vacuum tube?
(l) What is meant by a hard tube?
(m) What is a soft tube?
(n) What does a small blue glow appearing on the glass envelope of a hard-type tube indicate?
(o) What do we mean when we say that the plate current - plate voltage curve of the diode is linear over most of its range?

The Triode Tube

The development of the diode or two-element tube merely opened the door to the field of electronics. The diode tube has only limited applications. It can be used as a rectifier to change ac to dc, and in some other special circuits, but it cannot be used to amplify.

Early in the start of the twentieth century, the triode tube was developed. This tube touched off the rapid development in electronics that has occurred since then. The development of the triode tube led to the development of other multi-element tubes which have made possible circuits undreamed of not too many years ago.

You have already been introduced to the triode and know that the triode tube has three elements: a cathode, a grid, and a plate. The introduction of the grid between the cathode and the plate of the tube made it possible for a vacuum tube to amplify weak signals. You already have a general idea of how the grid works, but in this section we will review this idea, expand it, and then learn more about tube characteristics. It is extremely important for you to understand all the details of how a triode tube works, because the multi-element tubes that you will study later are simply triodes with additional elements added.

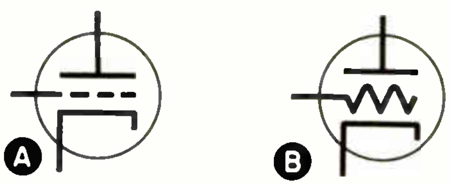


Fig. 7. Different ways of representing the grid on schematic diagrams.

HOW THE GRID WORKS

The third element inside the vacuum tube is placed between the cathode and plate and is called the grid. As the name implies, the grid is of open construction. The schematic symbol used to represent a grid is shown in Fig. 7A. On some old diagrams you might find the symbol shown in 7B used - this symbol is obsolete.

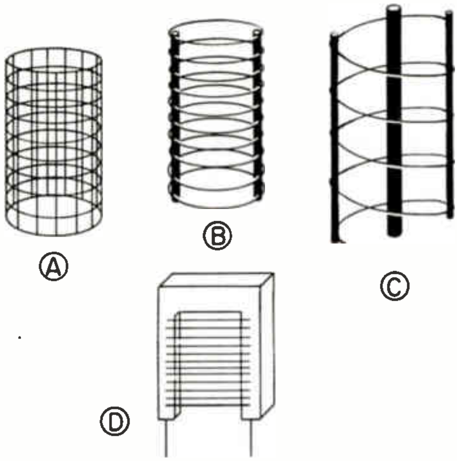


Fig. 8. Different types of grid structures.

Several different types of grid construction are shown in Fig. 8. Notice that in A the grid is made in the form of a spiral mesh, like a screen, whereas the grid shown in B is made up of a spiral-wound coil with the turns placed relatively close together. In C the same type of construction is used as in B, but the space between the turns is much greater. In D the grid is more or less rectangular in shape and is supported by the U-shaped elements at the end so that the grid is held in a very rigid position. The grid is supported by the frame and this type of construction is often referred to as a frame grid. It has this advantage over the other types: the grid wires can be placed very close together and very close to the cathode, which as you will see later makes it possible to make a tube with a much higher gain than that of the other types of grid.

Before we review how the tube amplifies, let's consider the effect of the grid on the flow of plate current when different voltages are applied to it. The first case we will take up is where the grid is connected directly to the cathode so that the voltage applied to it is zero.

Zero Grid Voltage.

When the grid is connected directly to the cathode we have the arrangement shown in Fig. 9. Here the plate is connected to the positive side of the B battery and the cathode is connected to the negative side of the battery. The grid is connected directly to the cathode, and a small battery is used to provide the heater voltage required to heat the heater, which in turn heats the cathode of the tube.

When the cathode is heated, it will emit electrons and they will fly off into the space surrounding the cathode. These electrons will form a cloud of electrons around the cathode. This cloud of electrons is called a space charge. Some of the electrons in the space charge will fall back to the cathode, others will be attracted by the positive potential applied to the plate of the tube and they will be drawn through the grid wires to the plate. A few electrons in travelling from the cathode

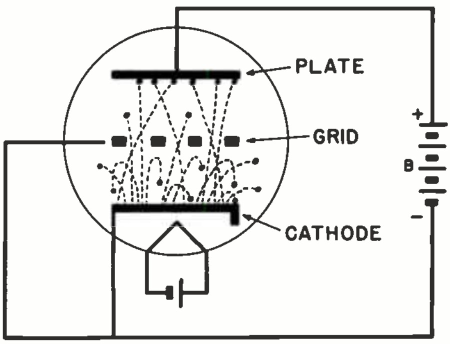


Fig. 9. When there is no grid bias, an average number of electrons flow to the plate, and the rest form a space charge between the cathode and the grid.

to the plate of the tube may accidentally strike the grid wires. These electrons will flow through the external circuit from the grid back to the cathode of the tube. As long as the grid is connected directly to the cathode, the tube acts very much like a diode and the grid has very little effect on the flow of plate current. The amount of plate current flowing will depend primarily upon the voltage applied between the plate and cathode of the tube and the spacing between the plate and cathode.

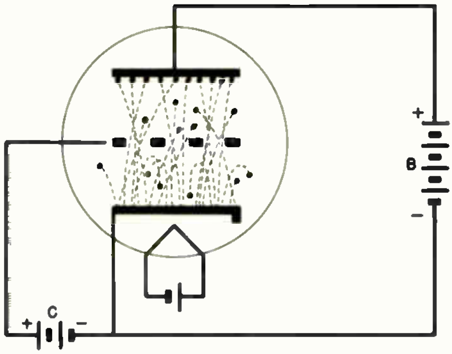


Fig. 10. Making the grid positive greatly increases the number of electrons moving to the plate.

Positive Grid Voltage.

Now if we modify the circuit shown in Fig. 9, by adding a small C battery in the grid circuit as shown in Fig. 10, we will have a positive voltage on the grid of the tube. This means that the grid will be slightly positive with respect to the cathode. The B battery used between the plate and cathode has a much higher voltage than the C battery, and therefore the plate will have a much higher positive potential than the grid.

With the C battery connected in the grid circuit, the positive voltage on the grid of the tube will attract electrons from the space charge around the cathode and start these electrons speeding toward the grid. By the time the electrons travel the short distance from the space charge to the grid most of them will be travelling at such a high speed that they will pass right through the grid wires and then come under the influence of the high positive voltage on the plate of the tube. Most of the electrons will therefore continue travelling towards the plate of the tube until they eventually will reach and strike the plate.

The number of electrons reaching the plate will be much higher than it was in the preceding case where there was no voltage applied to the grid. The grid is able to increase the number of electrons flowing to the plate because the grid is placed very close to the cathode, and even though there is only a low positive voltage applied to the grid it is able to pull many electrons from the electron cloud and start them on their way to the plate.

Because the grid has a positive voltage applied to it, there will be more electrons striking the grid than there were in the preceding case when there was no voltage applied to the grid. However, even though a few electrons will be attracted to the grid, a small positive voltage applied to the grid will increase the flow of plate current. If the positive voltage is made higher, in other words if the grid is made more positive, it will start to attract more and more electrons. Eventually a point will be reached where the grid will be taking many of the electrons that would normally flow over to the plate of the tube. Then, instead of causing the plate current to increase, the large number of electrons flowing to the grid of the tube will starve the plate so that the plate current will be less than it would be if the grid were operated at zero potential.

When the number of electrons striking the grid becomes high, the energy that these electrons give to the grid upon striking it may cause the grid to become very hot. As a matter of fact, if enough electrons strike the grid, the grid will become red hot. Keep this in mind. If you see a vacuum tube where one of the grids is showing a red heat, the number of electrons reaching the grid is too high--there is something wrong in the circuit.

Negative Grid Voltage.

If instead of putting a positive voltage on the tube grid, you put a negative voltage on it, you will have the circuit shown in Fig. 11. Now the negative potential on the grid of the tube repels the electrons coming from the space charge and drives them back to the space charge so that the number of electrons getting through the grid and reaching the plate of the tube is greatly reduced. As a matter of fact, if the negative voltage applied to the grid of the tube is made high enough, all electron movement between the cathode and the plate will be stopped --there will be no flow of electrons from the cathode to the plate of the tube.

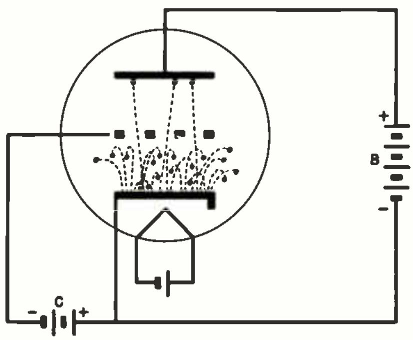


Fig. 11. Making the grid negative reduces the number of electrons moving to the plate.

Amplification Factor.

Current is actually a movement of electrons, and since the grid controls the electrons flowing from the cathode to the plate of the tube, the grid can control the current flowing from the cathode to the plate. The current flowing in the plate circuit is called the plate current. Changing the plate voltage on a triode tube will cause the plate current to change, but because the grid is closer to the cathode than the plate, it exerts a greater effect on plate current than the plate does. As a matter of fact, we may have to change the plate voltage on a tube as much as 100 volts to get the same change in plate current that can be obtained by changing the grid voltage only 1 volt. The exact ratio between the change in plate voltage and the change in grid voltage needed to get the same change in plate current is called the amplification factor. If we have to change the plate voltage 50 volts to get the same change in plate current we can get by changing the grid voltage 1 volt, the amplification factor is $\frac{50}{1} = 50$. A triode tube may have an amplification factor somewhere between 5 and about 100. The amplification factor of a tube depends primarily upon the ratio of the plate-cathode spacing to the grid-cathode spacing.

The amplification factor of a triode tube is a very important characteristic of the tube. It is a good indication of how much voltage gain you can expect to obtain from an amplifier stage using the tube. The total amplified voltage produced by the tube is equal to the amplification factor times the signal voltage applied between the grid and the cathode of the tube. It is not possible to get all of this amplified voltage out of the tube, because the tube has internal resistance, and part of the voltage will be dropped across this resistance, but in general the higher the amplification factor of a tube, the greater the gain we can expect to obtain from the tube.

To provide a short form for expressing the amplification factor of a tube, the Greek letter, mu, which is pronounced "mew", and written μ, is used as a symbol to represent the amplification factor. The amplification factor is often referred to as the mu of the tube. Thus a high-mu tube is a tube with a high amplification factor.

HOW A TRIODE AMPLIFIES

You have already had a brief explanation of how a triode amplifies a signal, so this will be primarily a review. Be sure that you completely understand how the tube amplifies because vacuum tubes are the heart of electronic equipment and if you do not know how tubes work, you will not understand the equipment in which they are used.

First, let's consider what type of signal we may want to amplify. For example, let's assume that we have a sine-wave signal that represents a certain sound. The amplitude of the signal is extremely weak. In order to use the signal to produce sound, we must increase the strength of the signal. The usual procedure is to first build up the voltage of the signal. An amplifier stage used for

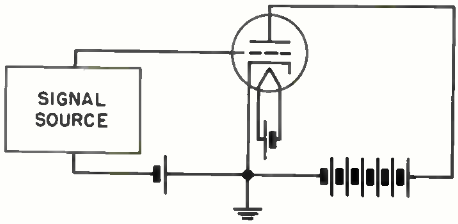


Fig. 12. The signal applied between the grid and the cathode will result in a varying plate current.

this purpose is called a voltage amplifier. If we apply the signal between the grid and the cathode of a vacuum tube connected as shown in Fig. 12, we know the ac signal in the grid circuit will produce variations in the plate current. We know this is true because the voltage applied to the grid has a pronounced effect on the flow of current to the plate. Thus the varying signal applied between the grid and the cathode of the tube will cause the plate current to vary.

However, we are interested in a varying output voltage. If we feed a weak signal voltage into the grid of a tube to amplify it, what we want is an amplified voltage in the output. The desired result can be obtained with a circuit like the one shown in Fig. 13. Now let's stop and consider what happens in this circuit first when the signal voltage is zero.

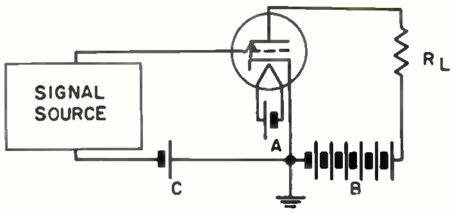


Fig. 13. The signal applied between the grid and the cathode will result in an amplified voltage in the output circuit.

When there is no ac signal applied to the input, a steady plate current will flow from the cathode of the tube to the plate of the tube. The exact amount of current is relatively unimportant, but it will depend upon the grid bias voltage supplied by the C battery, the plate voltage supplied by the B battery, the characteristics of the tube itself, and the size of the plate load resistor marked R_L.

The plate current flowing through the tube will flow from the negative side of the battery to the cathode of the tube. The heated cathode will give off electrons, which will flow through the space between the cathode and the grid of the tube over to the plate of the tube. The electrons will then flow from the plate of the tube through the plate load resistor R_L to the positive side of the battery, through the battery, back to the negative battery terminal. It sounds as though the electrons start from the negative terminal and then gradually flow around the circuit. Actually, the movement of electrons in the circuit is instantaneous. The electrons start moving in all parts of the circuit at the instant the tube is heated and power applied.

When current flows through the load resistor R_L, there will be a voltage drop produced across the resistor. The exact value will depend upon the current flow and the size of the resistor. We know this is true from Ohm's Law which states that $E = IR$. If there is a voltage drop across the dropping resistor, this means that the plate voltage, which is the voltage between the plate and the cathode of the tube, will be less than the B supply voltage by an amount equal to the voltage drop across the plate load resistor. Again, we know this is true because the sum of the voltage drops in a series circuit must be equal to the source voltage. In this series circuit we have the plate-cathode circuit of the tube in series with the load resistor. Therefore, the voltage across the tube equals the B supply voltage minus the voltage drop across the load resistor.

Before a signal is applied between the grid and the cathode of the tube the plate current will reach what is called a steady state. The voltage between the plate of the tube and the cathode will be some constant value.

Now let us consider what happens when a signal voltage like the sine-wave shown in Fig. 14A is applied between the grid and the cathode of the tube. Notice the first half cycle, which is numbered 1-2-3 on the input signal at A. This signal swings in a positive direction. When the ac input voltage swings positive, its polarity will be such that it will subtract from the grid bias voltage. This will make the grid voltage less negative. If we make the grid voltage less negative, we are moving it in a positive direction. When the grid becomes less negative it will allow more electrons to flow from the cathode of the tube to the plate. If the number of electrons moving in this part of the series circuit increases, then the total number of electrons in motion in all parts of

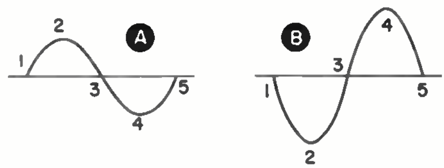


Fig. 14. A sine-wave signal voltage. The input signal is shown at A, and the output signal at B.

the series circuit must increase. The number of electrons flowing through the plate load resistor R_L increases. When this happens, the voltage drop across the plate load resistor increases, and there will be less of the B supply voltage between the plate and cathode of the tube. This means that the plate voltage will decrease.

The plate voltage will continue to decrease as the input signal moves from point 1 to point 2. The plate voltage will look like the wave shape shown in Fig. 14B and as the input moves from point 1 to point 2 on curve A, the output will move from point 1 to point 2 on curve B.

If the input signal begins to decrease from its peak point 2 back to zero voltage at point 3, the plate voltage between the plate and cathode of the tube will start to increase, because as the input signal moves from point 2 to point 3, the negative grid voltage on the tube will be increased. This means that there will be fewer electrons flowing from the cathode to the plate of the tube, and hence fewer electrons flowing through the load resistor. When the number of electrons flowing through this resistor decreases, the voltage drop across it decreases. Since the sum of the voltage drop across the load resistor and the voltage between the plate and cathode is always equal to the B supply voltage, the voltage between the plate and cathode of the tube must increase.

During the second half of the input cycle, when the input signal swings from point 3 to point 4, the signal is swinging in a negative direction, so it adds to the grid bias voltage applied to the tube by the C battery. This makes the grid even more negative, and reduces still further the number of electrons moving from the cathode to the plate of the tube. This reduction in electrons flowing in the series circuit means that there will be an even smaller voltage drop across the plate load resistor R_L and hence even more voltage between the plate and the cathode of the tube. As the input signal moves from point 3 to point 4 and then on to point 5, the output signal between the plate and cathode of the tube will follow the shape shown from point 3 to point 4 and then to point 5.

This variation in the voltage between the plate and cathode of the tube will be several times the amplitude of the input signal.

Before going ahead let's take a look at what we have in the plate circuit of the tube. Examine the waveform shown in Fig. 14B. Notice that a single ac cycle has been produced and the amplitude, frequency, and wave shape of the signal depend upon the signal applied between the grid and the cathode. This signal is an amplified version of the signal applied between the grid and the cathode of the tube.

The total signal applied between the grid and the cathode of the tube actually consists of a dc voltage applied by the C battery plus an ac voltage which is superimposed or added to it. The voltage between the cathode and grid will be the algebraic sum of the two voltages. When the signal voltage tends to swing the grid positive it will actually be subtracting from the C battery voltage so that the net negative voltage applied to the grid will be reduced. When the signal source tends to swing the grid in a negative direction, it will add to the C battery voltage so that the grid will be made more negative with respect to the cathode and will reduce the current flow through the tube.

The voltage in the plate circuit actually consists of a dc voltage applied to the tube with an ac voltage superimposed on it. Notice, however, that the ac voltage at the plate is inverted when compared to the input signal. In other words, when the input signal swings in a positive direction, the output signal swings in a negative direction, and when the input signal swings in a negative direction, the output signal swings in a positive direction. We say that the two signals are 180° out of phase.

Notice that the cathode of the tube in the circuit shown in Fig. 13 is grounded. This type of circuit is called a "grounded-cathode" amplifier. One of the characteristics of the grounded-cathode amplifier is that there is a 180° phase-shift between the input and output signals.

The signal voltage applied in series with the C bias voltage between the grid and cathode causes the plate current flowing from the cathode to the plate of the tube to vary. Actually, the current looks like a dc current with an ac current superimposed on it. Therefore, when we refer to the ac current in a tube, we mean the signal current or the variation in dc current caused by a signal applied between the grid and cathode. Remember that the current through a tube always flows from cathode to plate; it never reverses in direction. However, the variation in current through the tube due to the signal variations produces the same effect as we would have with a dc current with an ac current superimposed on it.

Remember that when you studied the amplification factor we pointed out that the amplification factor is based on a change in plate voltage and a change in grid voltage. Another way of expressing the same idea is in terms of ac voltages. We can say that it is the ratio of ac plate voltage to the amount of ac grid voltage required to produce the same ac plate current. As we mentioned previously, we are speaking of an ac current superimposed on a dc current.

SUMMARY

This section of your lesson is extremely important. Therefore we will not try to summarize it, but instead we suggest you go back and read over the entire section of the lesson again to be sure you understand it thoroughly. Read the section carefully and slowly and try to picture exactly what is going on inside the tube. Grasping clearly the idea of how a tube amplifies is extremely important. If you master this idea now, the remaining material in this book should be comparatively simple and you will also find that it will be much easier for you to understand how transistors amplify. When you are sure you understand this section of the lesson, test yourself by doing the following self-test questions.

SELF TEST QUESTIONS

(p) Why does the grid voltage have a greater effect on plate current than the plate voltage does?

(q) If the voltage applied between the grid and cathode of a triode tube places the grid positive with respect to the cathode, what effect will this have on the current flowing through the tube?

(r) If the grid is made negative

with respect to the cathode, and the voltage is slowly increased, what will happen to the plate current?

(s) What do we mean by amplification factor?

(t) Between what values would you expect the amplification factor of a triode to fall?

(u) When a signal voltage is amplified by a tube in a grounded cathode circuit, will the amplified signal be the same as the input signal?

(v) What do we mean by ac plate current?

(w) What is a grounded-cathode amplifier?

Tube Characteristics

Tube characteristics are important both to the engineer and to the technician. They are important to the engineer because he uses them in designing circuits. He must know the characteristics of the tube in order to select the correct value of parts to use along with the tube to get the best possible performance out of the circuit. They are important to the technician because often he will have to service equipment in which detailed service information is not available. By identifying the tube types used in the equipment and referring to a tube manual which describes the characteristics of the tubes used, he can obtain a great deal of useful information on how the equipment should work.

A great deal of information about tubes is given in the form of characteristic curves. The most important of these curves is the E_g–I_p (grid voltage-plate current) curve.

THE Eg-Ip CURVE

A tube is able to reproduce a signal only as long as a given change in grid voltage will produce a constant change in the plate current. Let's assume we have a tube with a voltage of 2 volts between the grid and cathode, and that the grid is 2 volts negative with respect to the cathode. If we increase this voltage from 2 volts to 2-1/2 volts, we know that the plate current will decrease. We could actually measure this change in plate

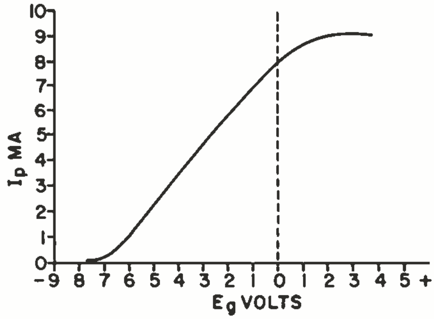


Fig. 15. A typical E_g-I_p characteristic curve.

current. Now, if we increase the voltage still further from 2-1/2 volts to 3 volts, we should get the same change in the plate current that we got when we changed the voltage from 2 volts to 2-1/2 volts. It is quite likely that we would get the same change, but eventually a point will be reached where changing the grid voltage 1/2 a volt will not produce this same change in plate current. When we reach this situation, we say that this tube becomes non-linear. This characteristic of a tube limits the amount of signal that can be applied between the grid and cathode of the tube.

The change in plate current that can be expected for a given change in grid voltage is shown on a curve called the E_g-I_p curve. A typical E_g-I_p curve is shown in Fig. 15. Notice how the plate current will be different for different values of grid voltage. Also notice that part of this curve is relatively straight but the two ends of the curve are quite bent. In the middle of the curve, we have a straight section that is called linear. This simply means that as long as the tube is operated within these limits a given change in grid voltage will produce the same change in plate current. As long as this change in plate current for a given change in the grid voltage remains constant, the output will be a faithful reproduction of the input signal. However, once the change becomes non-linear, the output will be distorted. This means the output signal will not be a faithful reproduction of the input signal.

Grid Bias.

The voltage applied by the C battery that is connected between the grid and cathode of a tube is called the grid bias voltage. To bias something means to fix or adjust it. The grid bias voltage fixes or adjusts the tube to operate under the most favorable conditions. The idea in back of a grid bias voltage is to apply a voltage to the grid that will place the tube operation on the desired portion of its characteristic curve. In the case of the amplifier we have been considering, we bias the tube to operate at the center of the linear or straight part of the characteristic curve. If you look at Fig. 15 you can see that the characteristic curve is comparatively straight between zero grid voltage and the grid voltage at -6 volts. Slightly to the right of zero grid voltage the curve becomes distorted, and to the left of -6 volts it also becomes distorted. Therefore, to get the tube operating on the linear part of the curve, we would apply a grid bias of -3 volts. With a grid bias of -3 volts, if the input signal drives the grid voltage 3 volts in a positive direction, we can expect the same change in plate current as we will get when the input voltage swings 3 volts in a negative direction. When the input voltage swings 3 volts in the positive direction, it will subtract from the grid bias, so the net voltage applied to the grid of the tube will be zero, and when it swings 3 volts in the negative direction, it will add to the grid bias so that the net voltage applied between the grid and the cathode of the tube will be -6 volts. Changing the grid voltage between these limits produces linear or constant changes in plate current.

On the other hand, if instead of a grid bias of -3 volts we had a grid bias of -5 volts and a 3-volt input signal, when the grid swings in the positive direction we will get a total voltage of -2 volts on the grid. This will produce a change in plate current. However, when the grid swings 3 volts in the opposite direction, due to the signal, we will get a total grid voltage of -8 volts. This will produce a change in plate current, but it will not be the same as the change produced when the signal went in the opposite direction. As a matter of fact, you can see from the curve that once the signal is past -7 volts little or no change in plate current will occur.

Under these circumstances, a plate current change that is different when the signal swings in one direction is different from the plate current change when the signal swings in the opposite direction and we get distortion. This means that the signal is not reproduced faithfully. For an input signal like the one shown in Fig. 16A we would get an output

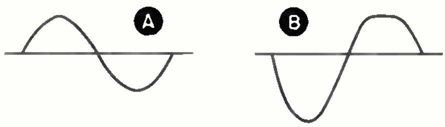


Fig. 16. The input voltage (A) and the output voltage (B) for a tube with excessive grid bias. Notice that the positive half of the cycle in the output is flattened.

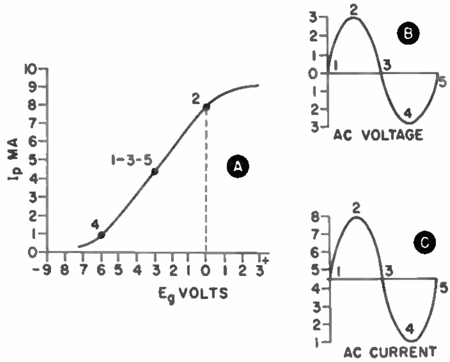


Fig. 17. A typical characteristic curve (A) showing how the grid voltage (B) and the plate current (C) vary on it.

signal like the one shown in Fig. 16B. Notice that in the first half cycle when the input signal swings in a positive direction the output signal swings in the opposite direction. However, the output signal is a faithful reproduction of the input signal. During the second half cycle, when the input signal swings in a negative direction, the output signal swings in a positive direction but is flattened off at the top. The amplified signal is not a faithful reproduction of the input signal. As we mentioned, this is called distortion, and this particular case of distortion is called amplitude distortion because the amplitude of one half of the cycle is compressed.

Now let us see how different values of grid bias affect the way in which a tube amplifies a signal. As you know, the grid voltage is made up of a fixed bias voltage, and the ac signal is superimposed on it. Assume we have a tube with the E_g-I_p curve shown in Fig. 17A. The center of the linear part of the curve is at -3 volts. Now suppose we have a bias voltage of -3 volts, and the signal voltage shown in Fig. 17B, which varies from 0 to 3 volts, is superimposed on it. When the ac signal is at the start of its cycle, point 1, it will be zero, so the total grid voltage will be -3, point 1 on the curve in 17A. When it increases to point 2, the total voltage will then be zero, point 2 on the curve. At point 3 the total voltage will be -3 volts again. At point 4 the total voltage will be -6, and at point 5 the total voltage will again be -3.

Now let's draw ourselves a picture showing what is happening to the plate current at each of these points. At point 1 the current is 4.5 ma; at point 2 it is 8 ma; at point 3 it is again 4.5 ma; at point 4 it is 1 ma; and at point 5 it is 4.5 ma again. This is shown in Fig. 17C.

As you can see, this is a faithful reproduction of the applied signal.

We often show these things all on one chart as shown in Fig. 18. Notice how we have drawn the signal voltage. By drawing it in this position, we can follow up to the point it represents on the characteristic

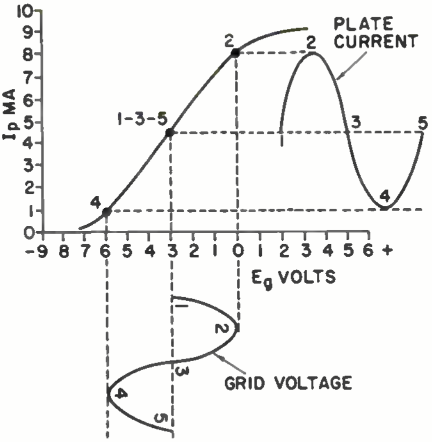


Fig. 18. We often show the information of Fig. 17. all on one diagram like this.

curve, and then show the curve for the corresponding plate current.

In Fig. 19 we have shown what happens when the operating bias is too high or too low. In Fig. 19A the operating bias is too high (too negative) and the result is that the grid is driven beyond the cut-off voltage

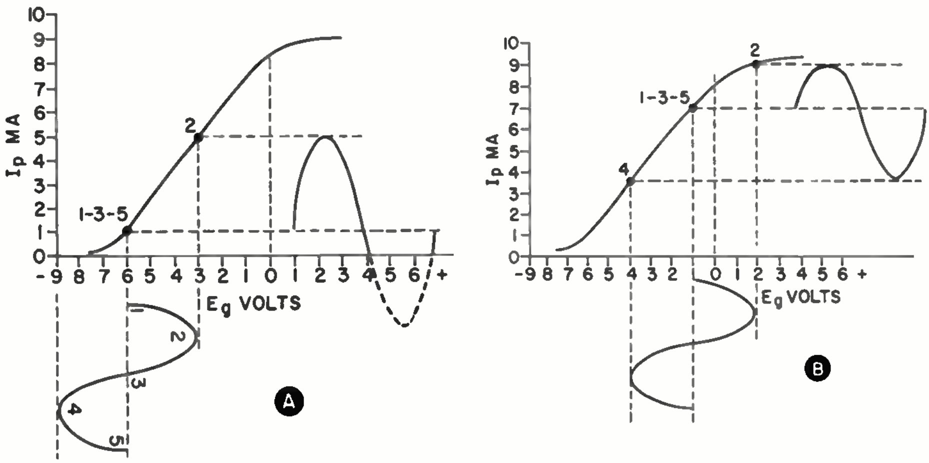


Fig. 19. The effect of too much bias is shown at A, and the effect of too little bias at B.

on the negative half of each cycle. This means that the plate current drops to zero and flattens off on the negative half of each cycle.

In Fig. 19B the operating voltage is not high enough, so during part of each cycle the grid becomes positive. When this happens the grid will start to take some of the electrons that would normally flow to the plate of the tube with the result that the top of the plate current curve is somewhat distorted. The variations in plate current for a given input signal are not linear when the bias is too low.

THE Ep-Ip CURVE

The characteristic curves shown in a tube manual are usually E_p-I_p (plate voltage-plate current) curves such as those shown in Fig. 20. Each curve is for a particular value of grid voltage. As you can see, when the plate voltage is higher it takes a higher grid bias voltage to cut off the flow of plate current. You can also get a good idea of how a change in grid voltage affects the plate current. Find the vertical line representing a plate voltage of 200 volts. Follow this line up until you see where it cuts the curve representing zero grid volts. It cuts this line just above the horizontal line representing a plate current of 6 ma. The -1 volt grid curve cuts the 200 volt line at about 3.5 ma. Thus, changing the grid voltage from zero to -1 volt will change the plate current from above 6 ma to 3.5 ma, or about 2.5 ma.

Notice how much more effective the grid is in controlling plate current than the plate is. We saw that a change in grid voltage of 1 volt caused a plate current change of 2.5

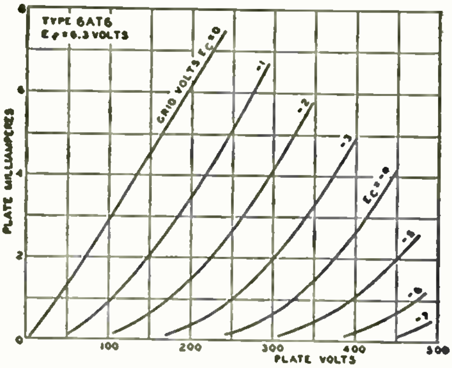


Fig. 20. E_p-I_p curves of the 6AT6 tube.

ma. The -1 volt grid curve cuts the 200 volt line at 3.5 ma. If we increase the plate voltage to 250 volts and keep the grid bias at -1 volt, the plate current will be 5 ma, which is a change of only 1.5 ma from the 200-volt point. Thus the change in plate voltage of 50 volts has less effect on the plate current than a change of grid voltage of only 1 volt.

PLATE RESISTANCE

As we mentioned previously, a tube has internal resistance. The tube, in amplifying a signal, acts very much like a generator with an internal resistance. This internal resistance of the tube is called the ac plate resistance. Because the tube has this internal plate resistance, the gain that can be obtained from a tube is never quite equal to the amplification factor of the tube.

The ac plate resistance is defined as the ratio of the change in plate volts to the change in plate current that it produces. The symbol r_p is usually used to represent the ac plate resistance of the tube. The symbol e_p is used to represent the change in plate voltage and the symbol i_p is used to represent the change in plate current. Thus the plate resistance of a tube in ohms is:

$$r_p = \frac{e_p}{i_p}$$

The ac plate resistance is one of the characteristics listed in manufacturers' tube manuals. When the tube manual gives a plate resistance, it is for a certain value of plate and grid voltage. The ac plate resistance of a tube is affected by the dc voltage applied to the plate and by the dc voltage applied to the grid. Therefore, if the tube is operated at voltages other than those listed, the plate resistance will not be the same as the values shown. In some engineering type tube manuals the plate resistance is given in the form of a curve so that the plate resistance can be determined with different values of plate and/or grid voltages.

So far we have been talking about the ac plate resistance of a tube. However, a tube also has a dc resistance. In a circuit such as the one shown in Fig. 13, with no input signal applied to the grid, a certain value of plate current will flow. This plate current produces a voltage drop across the tube and a voltage across the load resistor. You know from Ohm's Law that the resistance is equal to the voltage divided by the current, thus the dc resistance of the tube is equal to the dc plate voltage, which is the voltage between the plate and cathode of the tube divided by the dc plate current. Notice that both of these values are dc values and therefore we get a dc plate resistance for the tube. This particular characteristic is not usually of importance and when we refer to the plate resistance of the tube we will always mean the ac plate resistance unless we specifically say that we are referring to the dc plate resistance.

MUTUAL CONDUCTANCE

Another important tube characteristic is mutual conductance. The mutual conductance of the tube is usually represented by the symbol g_m. It is equal to the change in plate current divided by the change in grid voltage required to produce the change in plate current. The unit of mutual conductance is the mho. Notice that mho is simply ohm written backwards.

The formula for the mutual conductance of a tube is:

$$g_m = \frac{i_p}{e_g}$$

where g_m is the mutual conductance, i_p is the change in plate current and e_g is the change in grid voltage.

Since the change in plate current is usually in milliamperes, and the change in grid volts is in volts, the mutual conductance of the tube turns out to be a fraction of a mho. For example, if a change in grid voltage of 1 volt produces a change in plate current of 1 milliampere, the mutual conductance is:

$$g_m = \frac{.001}{1} = .001 \text{ mho}$$

To eliminate the fraction we usually express the mutual conductance of a tube in micromhos. A micromho is a millionth of a mho. Thus .001 mho equals 1000 micromhos.

RELATIONSHIP BETWEEN TUBE CHARACTERISTICS

So far we have discussed three important tube characteristics. They are the amplification factor, the plate resistance, and the mutual conductance. They can be represented by symbols as follows:

$$\text{Amplification factor} = \mu = \frac{e_p}{e_g}$$

$$\text{Plate resistance} = r_p = \frac{e_p}{i_p}$$

$$\text{Mutual Conductance} = g_m = \frac{i_p}{e_g}$$

Now let us look at how these characteristics are inter-related. If we multiply the plate resistance by the mutual conductance we get:

$$r_p \times g_m = \frac{e_p}{i_p} \times \frac{i_p}{e_g}$$

If you look at the second expression you will see that you have i_p both above and below the line so these two can be cancelled and we get:

$$\frac{e_p}{e_g}$$

We already know that e_p over e_g equals μ, which is the amplification factor. Thus the amplification factor is equal to the plate resistance times the mutual conductance or:

$$\mu = r_p \times g_m$$

EQUIVALENT CIRCUITS

It is sometimes difficult to visualize exactly what is happening inside a tube and in the circuit associated with the tube. However, by means of what is called an equivalent circuit we can analyze the performance of a tube more easily. A typical triode amplifier is shown in Fig. 21A and the equivalent circuit for this stage is shown in Fig. 21B. Notice that the tube is represented by a generator with a resistance connected in series with it. The voltage developed by the tube is equal to the amplification factor of the tube times the grid voltage which is written μe_g. We know, however, that the output voltage will be 180 degrees out of phase with the input, and we therefore put the minus sign in front of μe_g and indicate that the generator voltage is $-\mu e_g$. This indicates that the generator voltage, which is the amplified voltage produced by the tube, is 180 degrees out of phase with the grid voltage. In other words, at the instant the grid voltage swings positive, the generator will swing negative and then when the grid voltage swings negative, the generator voltage swings positive.

The ac plate resistance of the tube is represented by the resistor r_p. The voltage dropped across this resistance is lost as far as obtaining useful output is concerned.

Now let us see just exactly how

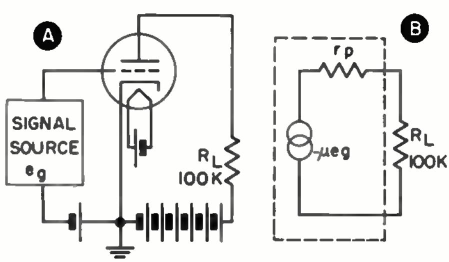


Fig. 21. A simple amplifier circuit is shown at A, and the equivalent circuit at B.

much gain we can obtain from the circuit shown in Fig. 21A. We could find the plate resistance of the tube and the amplification factor by looking them up in a tube manual. Let's assume that the amplification factor of the tube used in the circuit is 100, and the plate resistance of the tube is 50,000 ohms. The load resistance R_L is 100K ohms as shown on each diagram.

If the input signal e_g had an amplitude of 1 volt, then the generator output voltage $\mu e_g = 100 \times 1 = 100$ volts. The 100 volts produced by the generator divides between the 50,000-ohm plate resistance r_p and the 100,000 ohm load resistance, R_L. Thus, since the load resistance is twice the size of the plate resistance, the voltage across the load resistance will be twice the voltage across the plate resistance. This means that one-third of the 100 volts at the generator will appear across the plate resistance, and two-thirds, or approximately 66 volts, will appear across the load resistance. From this we can see that for an input voltage of 1 volt, the useful output voltage across the load resistor is 66 volts. The gain of the stage is equal to the output voltage divided by the input voltage, which in this case is $66 \div 1 = 66$. Therefore in this circuit using a tube with an amplification factor of 100, we obtained a gain of 66.

Consider what would happen if instead of using a 100,000-ohm plate load resistor we used a 200,000-ohm plate load resistor. Now the plate load resistor is four times the plate resistance and therefore there will be four times as much voltage across the load resistor as across the plate resistance. This means that one-fifth of the voltage, or 20 volts, will appear across the plate resistance, and four-fifths, or 80 volts, will appear across the load resistor. Therefore, the output voltage in this case would be 80 volts and the gain of the stage would be 80.

From this you might think that all we have to do is make the plate load resistor very large and we would get even more gain. This is true up to a point, but remember that the current flowing in the plate-cathode circuit of the tube must also flow through the plate load resistor. The current through the plate load resistor produces a voltage drop across the load resistor. The higher the resistance of the resistor the greater the voltage drop across it. Since the voltage available between the plate and the cathode of the tube is equal to the B supply voltage minus the voltage drop across the plate load resistor there will be very little voltage available between the plate and cathode of the tube if we make the plate load resistor too large. Eventually a point is reached where the plate-cathode voltage is so low that the amplification factor of the tube begins to fall off. When this point is reached, increasing the size of the plate load resistor results in no further increase in the gain of the stage.

A way to calculate the gain of a triode stage is by the equation:

$$\text{Stage gain} = \mu \times \frac{R_L}{R_L + r_p}$$

This equation shows that the larger the value of R_L with respect to r_p, the greater the gain of the stage. However, as we pointed out there are practical limits to this because the plate voltage does drop too low if R_L is made too large.

SUMMARY

In this section of this lesson you have greatly expanded your knowledge of tubes. There is a great deal of information in the section, and you should not expect to master and retain all the ideas presented with one reading. Be sure to go back and review this section of this lesson several times. This is particularly important because practically all of your later lessons are going to be based on the assumption that you understand how tubes work.

Characteristic curves are important because they give you an indication of how the plate current of a tube is going to change with changes in grid voltage. You can also see how large a signal a tube can handle without distortion. We do not expect you to remember what the characteristic curves we show look like, but you should remember that they are curved on both ends and this curvature in the characteristic curve limits the amount of signal that can be handled without distortion.

The three characteristics of tubes that you should remember are the plate resistance, the amplification factor, and the mutual conductance. Remember what they are:

$$\text{Plate resistance} = r_p = \frac{e_p}{i_p}$$

$$\text{Amplification factor} = \mu = \frac{e_p}{e_g}$$

$$\text{Mutual conductance} = g_m = \frac{i_p}{e_g}$$

You should also remember that the amplification factor of a tube is equal to the plate resistance times the mutual conductance.

Remember what the equivalent circuit of the simple triode amplifier looks like. We consider the tube as a generator with a resistance in series with it. The output voltage of the generator is equal to $-\mu e_g$. The resistance in series with the generator is equal to r_p, the plate resistance of the tube.

Remember that the equivalent tube circuit applies only to the ac signal-amplifying operation of the tube. We did not discuss the dc operating voltages when we discussed the equivalent circuit. Later you will see that the equivalent circuit is very valuable in analyzing and understanding the operation of amplifier stages. It will be particularly useful when we study how amplifiers amplify signals of different frequencies.

SELF-TEST QUESTIONS

(x) What is an E_g-I_p curve?

(y) What do we mean by the linear portion of the E_g-I_p characteristic curve?

(z) What is a grid bias voltage?

(aa) What will be the effect of operating the tube with too low a grid bias?

(ab) What happens when a tube is operated with too high a grid bias?

(ac) What is an E_p-I_p curve?

(ad) What is the ac plate resistance of a tube?

(ae) What is the dc plate resistance of a tube?

(af) What is the mutual conductance of a tube?

(ag) In what units is the mutual conductance of a tube measured?

(ah) What is the relationship between amplification factor, plate resistance and mutual conductance?

(ai) What do we mean by an equivalent circuit of an amplifier stage?

(aj) In an equivalent circuit used to analyze the performance of a triode amplifier, what is the generator voltage?

(ak) In the equivalent circuit of a triode amplifier, what is the resistance in series with the generator?

(al) If a triode amplifier has an amplification factor of 50 and a plate resistance of 10,000 ohms, what will the stage gain be when the amplifier is used in a circuit with a 90,000-ohm load resistor?

Multi-Element Tubes

From your study of the preceding section you can see that the triode tube is a very useful device. However, the triode has some definite limitations, and these limitations can be overcome by adding additional elements to the tubes. First, before studying the multi-element tubes, let's consider one of the big disadvantages of the triode in order to be able to understand the advantages of multi-element tubes better.

PLATE-TO-GRID CAPACITY

You will remember from your lesson on capacitors that a capacitor consists of two metal plates placed close to each other. In capacitors the metal plates are deliberately placed near each other in order to have capacity. However, when two pieces of metal that are insulated from each other are brought near each other, we will have a capacitor whether we want it or not; for example, the plate and grid of a vacuum tube. The plate, as you know, consists of a cylindrically shaped piece of metal. The grid is a spiral wire or mesh. The grid and plate, since they are placed near and are insulated from each other, actually form the two plates of a capacitor. Because the grid wire is small and there is a reasonable amount of space between the plate and grid, the capacity is small, but in some circuits, particularly where fairly high frequencies are involved, there is a high enough capacity to introduce a number of undesired effects.

As you know, at a high frequency even a small capacitor has a fairly low reactance. Therefore if a triode tube is used to amplify a high-frequency signal, the amplified signal present in the plate circuit of the tube can be fed back into the grid circuit through the plate-to-grid capacity. Under certain circumstances this signal fed from the plate back to the grid can be in phase with the signal applied to the grid of the tube so it will add to the input signal. This increase in the amplitude of the input signal in turn produces a still greater signal in the output.

The increased signal in the output in turn increases the signal fed back into the grid circuit which makes the grid signal still stronger. This increases the output signal still more. This action goes on and on until eventually a point is reached where the signal in the plate circuit takes control in the grid and the tube begins to generate its own signal voltage. This is called oscillation. The signal fed from the plate of the tube back to the grid is called a feedback signal. When the feedback signal, or more simply the feedback, is of the correct phase so that it adds to the grid signal, the tube will oscillate, which means generate its own signal, and hence will not work as an amplifier.

Not all triode amplifier stages will oscillate. It is possible to overcome this problem by means of suitable circuitry. However, in most applications it is easier and more practical to overcome the problem in the tube itself. This is accomplished by the addition of another grid inside the tube. The second grid is placed between the plate and the first grid to shield or screen the grid from the plate and is called the screen grid or simply the screen. The screen is also called the number-two grid.

THE SCREEN-GRID TUBE

The screen-grid tube is a four-element tube usually called a tetrode. The four elements are the cathode, the grid, the screen grid, and the plate. A top view of a screen-grid tube showing the four elements is shown in Fig. 22A, and the schematic symbol for the screen-grid tube is shown in Fig. 22B.

The screen grid is usually made

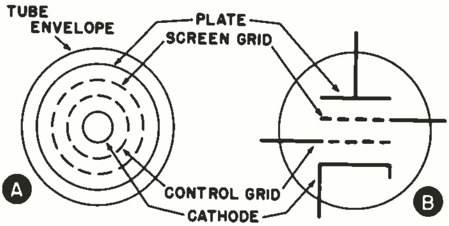


Fig. 22. A screen-grid tube: A, a cross section showing the arrangement of the elements; B, the schematic symbol.

just like a spiral-wound coil. The screen-grid wires are placed directly behind the grid wires so they are completely hidden from the electrons by the grid wire. A cut-away view of the tube would look something like Fig. 23. Notice that the screen-grid wires are immediately behind the control-grid wires.

The effect of the screen grid is to break the capacity between the plate and the control grid into two separate capacitors. In other words, there is a small capacity between the plate of the tube and the screen, and another capacity between the screen of the tube and the grid. These two capacitors are connected in series so that the net plate-to-grid

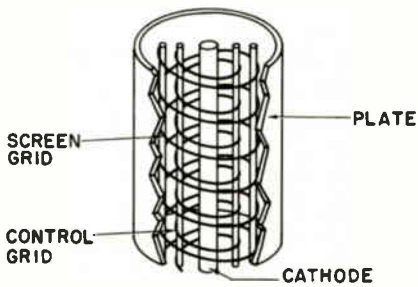


Fig. 23. A cut-away view of a screen-grid tube.

capacity is much smaller than it is in the triode. In addition, the screen grid of a tube is normally operated at signal ground potential. Therefore any energy fed from the plate of the tube back towards the grid is fed to ground at the screen. This practically eliminates the feedback from the plate to the grid of the tube.

The dc potential applied to the screen is always positive with respect to the cathode. Usually the voltage placed on the screen grid is about half the voltage applied to the plate of the tube, but in some tubes the screen-grid voltage may be as high as the plate voltage. This may appear to be a contradiction because we said that the screen grid is operated at the signal ground potential. Actually, it is possible to have a tube element at signal ground potential and still have a positive or negative dc voltage applied to it. You will remember that a capacitor offers a low reactance to the flow of ac through it. Therefore we can ground the screen of a tube so far as signal is concerned by putting a capacitor between the screen of the tube and ground. If the capacitor is large enough, its reactance to the signal is so low that the screen is practically at signal ground potential. At the same time, since a capacitor does not permit the flow of dc through it, you can connect it to a dc voltage source and apply a positive voltage to the screen.

The schematic diagram of a typical screen-grid amplifier circuit is shown in Fig. 24. Here the input signal is fed between the grid and the cathode of the tube as in the case of the triode. Electrons flow from the cathode of the tube to the plate of the tube as in the triode. However, in the tetrode tube the number of electrons flowing will depend upon the plate voltage, the grid voltage and the screen-grid voltage. In fact, the number of electrons flowing from the cathode towards the plate will depend much more on the screen and grid voltages than on the plate voltage.

As in the triode tube, small changes in grid voltage produced by the ac signal applied to the input will produce comparatively large changes in plate current. These

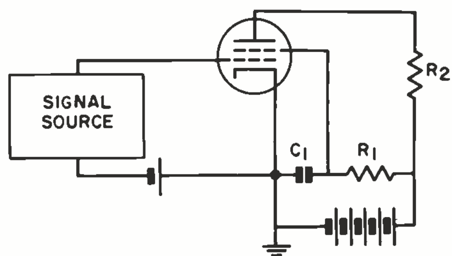


Fig. 24. Schematic of a screen-grid amplifier. R1 is used to reduce the B supply voltage for the screen, and is called the screen voltage dropping resistor, or simply the screen dropping resistor. C1 is the screen by-pass capacitor. R2 is the plate load resistor. The heater of the tube has been omitted to simplify the diagram.

changes in plate current will cause the voltage between the plate and cathode of the tube to vary, and this voltage will be the amplified output signal.

A few of the electrons flowing from the cathode of the tube towards the plate of the tube will strike the screen. Thus there will be a small current flowing from the cathode of the tube, to the screen grid, through the power supply and back to the cathode.

Because the plate voltage has

much less effect on the plate current in a screen-grid tube than in a triode tube, screen-grid tubes have a much higher amplification factor than triodes. In addition they have a much higher plate resistance.

Disadvantages.

Although the tetrode is a great improvement over the triode, it does have certain disadvantages. One of these disadvantages, which led to the development of the pentode or five-element tube, occurs because of secondary emission.

As you already know, electrons travelling from the cathode of the tube to the plate reach a fairly high speed, and when they strike the plate they may knock other electrons off the plate. In a diode and a triode the plate is the only positive element in the vicinity of these loose electrons, and therefore they are attracted back to the plate. However, in a tetrode we have, in addition to the plate, the screen grid with a positive voltage applied to it. If the plate voltage is substantially higher than the screen voltage, the electrons will be attracted back to the plate, but if the screen voltage is almost equal to or even higher than the plate voltage, then the electrons knocked off the plate of the tube will be attracted to the screen instead of to the plate. Thus, for every electron reaching the plate from the cathode there may be two or three electrons emitted by the plate. This means that if the grid swings in a positive direction more electrons will strike the plate, knocking still more electrons off the plate. The net result can be that the plate current decreases when the grid swings positive. To prevent this undesirable action a third grid called a suppressor grid was added to develop the pentode tube.

PENTODE TUBES

The pentode tube is a tube with five elements. It is simply a refinement of the screen-grid tube, which is made by adding an additional grid between the screen grid and the plate of the tube. Thus, the pentode has three grids. The third grid, or suppressor grid, gets its name because it is put in the tube to suppress or eliminate the undesirable effects of secondary emission occurring at the plate of the tetrode tube. Because it is also the third grid it is sometimes called the number-three grid.

The suppressor is usually connected directly to the cathode of the tube, but sometimes it is connected directly to ground. It has very little effect upon the electrons travelling from the cathode of the tube towards the plate. These electrons are attracted from the cathode by the positive potential on the screen of the tube. They are accelerated by this voltage, but as they approach the screen, instead of stopping at the screen they pass right through the screen wires. Once they get through these wires they are attracted to the plate by its positive voltage, which is usually higher than the voltage on the screen. In the pentode tube, after the electrons have passed the screen they are moving at a fairly high velocity, and they travel right on through the suppressor grid to the plate. The suppressor does not have any appreciable effect on the progress of the electron as it moves from the cathode toward the plate.

When the electrons strike the plate they knock other electrons off the plate, as in the case of the screen-grid tube. However, these electrons are travelling at a comparatively low speed when they are knocked off the plate. The suppressor, which is connected to the cathode or to ground or may be operated with a low negative voltage applied to it, repels the electrons emitted from the plate by secondary emission, and these electrons move back to the plate of the tube. Thus, the addition of the suppressor grid eliminates the undesirable current flow from the plate of the tube to the screen grid, which we found could occur in the tetrode tube.

Because the pentode tube has a screen grid, it has the low plate-to-grid capacity that we found in the tetrode tube. In addition, the pentode has the advantage that it does not suffer from the adverse effects of secondary emission.

It might be well to point out now that the plate voltage on a tetrode or a pentode tube has very little effect on the plate current. The voltage on these tubes can be varied over wide limits without appreciably changing the plate current that will flow in the tube. For a given grid voltage, the plate current will depend primarily upon the screen voltage. The screen is the electrode that starts the electrons moving from the cathode to the plate and has more effect on the number of electrons that will flow from the cathode to the plate than the plate voltage does.

You will remember that one of the characteristics of vacuum tubes that we studied was the amplification factor. The amplification factor is the ratio of the plate voltage required to produce a given change in plate current to the grid voltage required to produce the same change in plate current. Remember that the formula is:

$$\mu = \frac{e_p}{e_g}$$

Since the plate voltage has very little effect on the plate current in a tetrode or pentode tube, it takes a very large change in plate voltage to produce the same change in plate current that can be produced by a small change in grid voltage. Thus, the amplification factor of a tetrode or a pentode tube is very high. We mentioned that the amplification factor of a triode may be somewhere between about 5 and 100. A triode with an amplification factor of 100 is called a high-mu triode because it has a high amplification factor for a triode. However, pentodes with amplification factors of over 1000 are common.

BEAM POWER TUBES

Another tube that solves the problem of secondary emission in screen-grid tubes is the beam power tube. The beam power tube, like the screen-grid tube, is a tetrode or four-element tube in which this problem has been overcome. A sketch of a beam power tube is shown in Fig. 25.

Notice the shape of the cathode of the tube. It has two flat surfaces. The construction of the cathode is such that most of the electrons will be emitted by the flat surfaces and hence there is a tendency for the electrons to form into two beams, one on each side of the cathode. Only one of these beams is shown in the drawing.

Notice the two small additional plates between the screen grid and the plate of the tube. These plates are called the beam-confining or

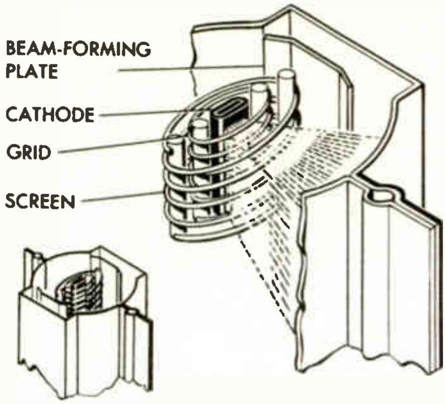


Fig. 25. A beam-power tube.

beam-forming plates. The beam-forming plates are connected internally to the cathode of the tube and repel electrons. These plates act to keep the electrons concentrated into the two beams that are formed at the cathode. The electrons flowing from the cathode of the tube towards the plate are bunched together in these beams.

You will remember that an electron has a negative charge and will repel other electrons. Therefore, if an electron travelling at a high speed from the cathode of the tube to the plate knocks additional electrons off the plate, these loose electrons, which will be travelling at a low speed, will be repelled back to the plate by the negative charge on the electron beam.

Beam power tubes are used in radio, TV, and industrial electronic equipment. Also, large beam power tubes are used in transmitters. The beam power tube has proven superior to the pentode in applications where large amounts of power must be handled.

OTHER TUBE TYPES

The diode, triode, tetrode, and pentode are the basic tube types. However, there are many special tubes that have been manufactured for special applications. In addition, there are tubes that are simply combinations of several tubes in the same envelope.

An example of a tube where several types have been combined in one envelope is the tube with a triode and two diodes in it. A schematic of this type of tube is shown in Fig. 26. This tube is called a duo-diode-triode. As this name indicates, it has two diodes and a triode in one envelope.

Other tubes consist of two triodes in one envelope such as shown in Fig. 27. In the schematic shown at A, a common cathode is used for the two triode sections; in the one shown at B, each triode section has its own separate cathode.

Another combination type is the triode-pentode shown in Fig. 28. This type of tube is used in radio, TV, and many industrial applications. In this way, one tube can be made to do the work of several.

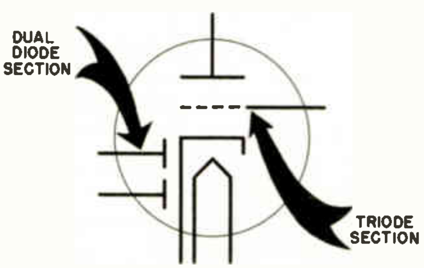


Fig. 26. A combination tube containing two diodes and a triode. This is called a duo-diode triode.

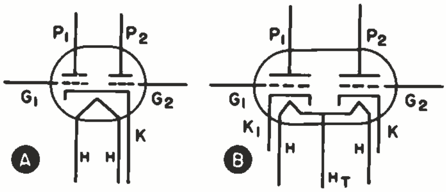


Fig. 27. Dual triode tubes. The tube at A has a common cathode. The tube at B has two separate cathodes. Notice this tube has a single heater with a center tap. Not all tubes of this type have tapped heaters.

The pentagrid tube, shown in Fig. 29, is an interesting tube. It is called a pentagrid tube because it has five grids. Tubes of this type are used as combination mixer-oscillators in radio and TV receivers. We will study this type in more detail later.

You will also run into a tube called the Compactron. The Compactron is simply a tube where several types are combined in one envelope. Compactron tubes have a base with twelve pins so it is possible to put a number of tubes in the same envelope. Some Compactrons may have three triodes in the one envelope; some may have two pentodes in the same envelope. Others are combinations of diodes, pentodes and triodes. Tubes such as

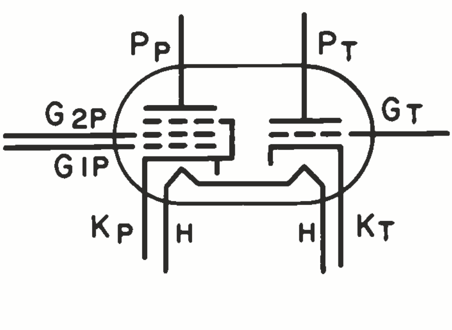


Fig. 28. A pentode-triode.

Compactrons offer an advantage over tubes in separate envelopes inasmuch as they are somewhat less expensive to manufacture than separate tubes, and also it is possible to make more compact equipment because the space occupied by the multi-function tube is considerably smaller than the space that would be required for separate tubes.

The Nuvistor tube is another important tube. Most Nuvistors are simply triode tubes but they are very small tubes. They look almost like a flat thimble. The advantage of the Nuvistor tube, in addition to its small size, is the very low capacity between the elements and the compara-

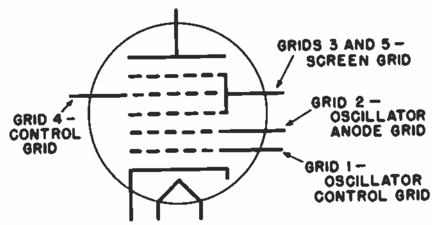


Fig. 29. A pentagrid tube.

tively high gain that can be obtained with this type of tube. Because of the very small size of the elements used in the Nuvistor tube, this tube is quite widely used as an rf amplifier in TV receivers. Later, when we start studying specific amplifier circuits, you will see why such a tube as the Nuvistor would be advantageous as an amplifier at very high frequencies.

While not all Nuvistor-type tubes are triodes, almost all the ones you are likely to encounter will be triodes. There have been a few tetrode Nuvistors manufactured, but these were primarily for industrial applications rather than entertainment-type equipment. In addition, due to manufacturing difficulties and competition from transistors, Nuvistors other than the triode types have not been too widely used.

CHARACTERISTIC CURVES

The characteristic curves for pentode tubes are quite different from those for triodes. The E_p-I_p curves for a typical pentode tube are shown in Fig. 30. Notice that the curves bend rather sharply at the left, but then are quite flat. For example, notice the curve representing the plate current with a grid bias of -2 volts. As we start at the lower left of the graph this curve rises rapidly until we reach the line representing the plate voltage of 50 volts. At this point the plate current is almost 2 milliamperes. Then, if the voltage on the plate of the tube is increased from 50 volts on up to 450 volts, there is very little change in the plate current. This simply demonstrates what we mentioned previously; in a pentode tube the plate voltage has very little effect on the current-the plate current is primarily determined by the grid voltage and the screen voltage.

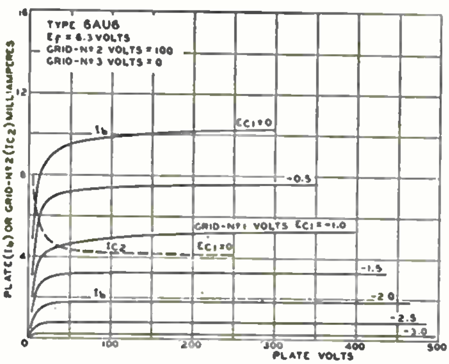


Fig. 30. E_p-I_p curves for the 6AU6 tube.

The characteristic curves for beam power tubes are quite similar to those for pentode tubes. We have not shown curves for screen-grid tubes, because the screen-grid tube has been replaced by the pentode and beam power tubes in modern design.

SUMMARY

In this section of the lesson you have studied multi-element tubes. Remember that the screen-grid tube was developed to eliminate the undesirable effects produced by the high plate-to-grid capacity found in triode tubes. A screen-grid tube is a tube with a high amplification factor and low plate-to-grid capacity. However, because the screen is operated with a positive voltage, the electrons emitted from the plate by secondary emission were attracted both to the plate and to the screen of the tube. Some of the electrons would flow to the screen and introduce a number of undesirable effects.

To overcome the problem created by secondary emission in the screen grid tube, the pentode or five-element tube was developed. The pentode contains three grids: a control grid, a screen grid, and a suppressor grid. These grids are often referred to as grids 1, 2, and 3.

The undesirable effects of secondary emission can also be overcome by constructing a flat type of cathode and using beam-forming plates. This led to the development of the beam power tube. The beam power tube is a tetrode tube, but it differs substantially in performance and construction from the old screen-grid tube.

In addition to the screen grid, pentode, and beam power tubes, there are a number of special tubes. These tubes may actually be tubes designed for one specific application such as the pentagrid converter type tube, or may simply be two or three different tubes combined in the one envelope. Tubes of this type consist of twin triodes, triode-pentode combinations, and duo-diode-triodes.

SELF-TEST QUESTIONS

(am) What is the purpose of the screen grid in a tetrode tube?

(an) What dc operating potential is applied to the screen of a tube?

(ao) How is the screen of a tetrode tube operated in order to provide maximum shielding between plate and grid?

(ap) Name the five elements in a pentode tube.

(aq) What elements control the flow of plate current in the pentode tube?

(ar) How are the undesirable effects of secondary emission overcome in the beam power tube?

(as) What is a Compactron?

(at) What is a Nuvistor?

Special Tube Types

There are a number of special-purpose tubes that have been designed for particular applications. Many of these tubes were designed for use in transmitting and industrial electronic equipment, but some of these tubes will also be found in equipment that the radio-TV serviceman may be called on to service. In most cases it is not particularly difficult to understand how these tubes operate.

GAS-FILLED DIODES

We mentioned earlier that gas is sometimes deliberately introduced into a tube. A common application of this principle is found in some diodes where mercury is placed inside the tube. The mercury vaporizes, filling the inside of the tube with mercury vapor. This particular type of tube makes an excellent rectifier. One of its characteristics is that the voltage drop across the tube is almost constant regardless of the current flowing through the tube. Let us look into the mercury-vapor diode tube and see how it differs from a high-vacuum diode.

The circuit shown in Fig. 31 can be used to compare the characteristics of the vacuum diode and the mercury-vapor diode. If we use a vacuum diode in this circuit, we will find that when the voltage applied to the circuit is zero, the voltage between the plate and the filament of the tube is zero, and the current flowing in the circuit is zero. If we increase the voltage to 5 volts, we

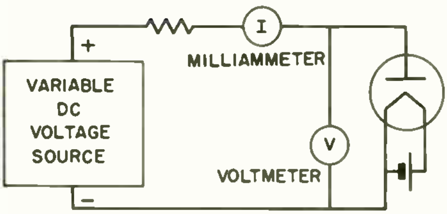


Fig. 31. A circuit for comparing vacuum and gas-filled diodes.

will notice a reading on the voltmeter indicating that there is voltage between the plate and the filament of the tube and that current is starting to flow in the circuit. If we then increase the variable voltage source to 10 volts, we will find that the voltage between the plate and the filament of the tube has increased and that the current flowing in the circuit has increased. If we keep increasing the voltage in 5-volt steps, we will find that the reading on the voltmeter will gradually increase, and at the same time the current flowing in the circuit will increase. This will continue in this way until a plate voltage is reached where all the electrons being emitted by the tube are being attracted to the plate. Then, increasing the voltage would result in little or no increase in current flowing through the tube. If we plotted the curve to show the relationship between the voltage across the tube and the current flowing through the tube, we would get a curve like the one shown in Fig. 32A.

If we perform the same experiment with a mercury vapor diode, we would obtain somewhat different results. If we started with the voltage from the variable voltage source at zero, we would find the reading on the voltmeter connected between the plate and filament of the tube was zero and again that the current flowing in the circuit was zero. When we increase the source voltage until we get a reading on the voltmeter of about 5 volts, there will be a current flowing in the circuit. A further increase to 10 volts would result in a further increase in the current flowing in the circuit. If we increase the voltage to 15 volts, we would get another increase in the current flowing in the circuit. However, when we increase the voltage above 15 volts something unusual will happen. At some voltage above 15 volts, the tube will fire; this means that it will suddenly start to glow with a blue glow. When this happens, the reading on the voltmeter connected between the plate and the filament of the tube will drop back to 15 volts. We would find that increasing the voltage from the variable voltage source further would not result in any increase in the voltage between the plate and the filament of the tube. However, the current flowing in the circuit would continue to increase as long as we increased the voltage of the source. If we plotted a curve to show the relationship between the voltmeter reading and the milliammeter readings for this tube, we

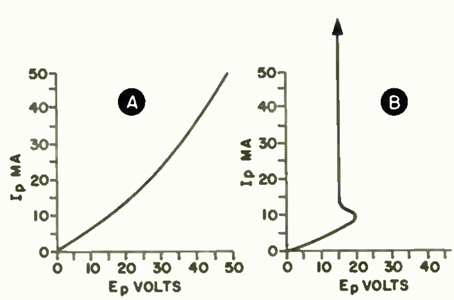


Fig. 32. The E_p-I_p curve for a vacuum diode is shown at A, and the curve for a mercury-vapor diode is shown at B.

would get a curve similar to Fig. 32B. Notice that up to some voltage slightly over 15 volts this curve is very similar to the curve obtained for the vacuum diode, but when the firing point is reached, the voltage between the plate and the filament of the tube drops down to 15 volts and then remains constant at this value, while at the same time the current through the tube can increase almost indefinitely.

Now let us see what is happening inside the mercury vapor tube. When the low positive voltage is applied to the plate of the tube, electrons emitted by the filament are attracted to the plate. These electrons do not reach any great velocity, so they simply travel over to the plate of the tube. Some of them will strike the gas molecules inside the tube, but they are not traveling at a high enough speed to knock any electrons off the molecules. As the plate voltage is increased, eventually a voltage is reached where the electrons travelling from the filament to the plate of the tube reach a high enough speed to knock electrons off the gas molecules that they strike. When this happens, the electrons knocked off the gas molecules travel over to the plate of the tube, thus increasing the number of electrons reaching the plate. At the same time, the molecules that have had some electrons removed will have a positive charge on them, and they will travel over towards the filament of the tube. As these molecules enter into the area of the space charge around the tube filament, they will pick up electrons from the space charge, lose their positive charge, and then begin to drift away from the tube cathode. They, in turn, will be hit by other electrons travelling from the filament of the tube to the plate, which will knock electrons off the gas molecule again, give it a positive charge, and once more it will start back to the space charge to pick up additional electrons in order to get rid of the positive charge.

If the right amount of mercury vapor is present in the tube, the gas molecules will neutralize or eliminate the effects of the space charge around the filament of the tube. Thus, the electrons will be able to leave the filament and travel directly to the plate of the tube with little or no opposition. This results in a tube with a very low internal resistance. The voltage drop across a tube of this type is almost constant and is about 15 volts regardless of the current flowing through the tube.

When the gas inside the tube ionizes, it gives off a bright blue glow. When this happens we say that the tube fires. The firing point for a mercury vapor tube is slightly above 15 volts, but once the tube has fired the voltage drop across the tube will drop back to approximately 15 volts and remain essentially constant at this value.

Mercury vapor diodes are used as rectifiers. It is a great advantage to have rectifier tubes that have a constant voltage drop, particularly in equipment where the current drawn from the power supply varies appreciably. If the current varies and a vacuum type rectifier is used, there will be considerable variation in the voltage drop across the tube, and as a result a variation in the output voltage from the power supply. However, in a mercury vapor tube, since the voltage drop across the tube is practically constant regardless of the current flowing through it, the output voltage from

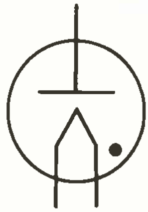


Fig. 33. Dot inside the envelope is used to indicate a gas-filled tube. This is the schematic symbol for a gas-filled diode.

the power supply will be almost constant even though the current drawn from the supply and through the tube may vary appreciably.

To distinguish tubes containing gas from vacuum tubes in a schematic diagram, a dot is usually placed inside the tube envelope on the schematic symbol. The schematic symbol for a gas-filled diode rectifier is shown in Fig. 33.

Mercury is not the only gas used in gas-filled tubes, but it is the most commonly used in rectifier tubes. Mercury-vapor tubes are not found in modern radio and TV receivers but are used in many other applications.

An example of another gas-filled diode is the Tungar tube. This tube is somewhat similar to and operates

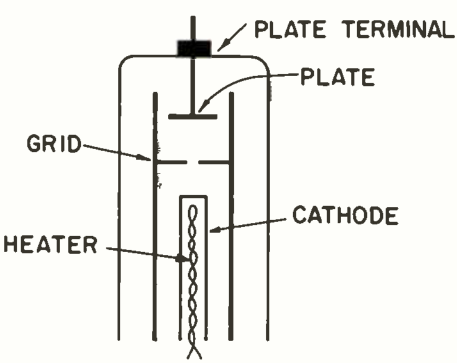


Fig. 34. Construction of a thyratron.

on the same principle as the mercury-vapor diode, but instead of mercury, argon gas is introduced into the tube. This tube has many of the characteristics of the mercury-vapor diode; it has an almost constant voltage drop of about 10 volts across it. It is particularly suitable for applications requiring a low voltage and a high current.

THE THYRATRON

In addition to mercury-vapor diodes, mercury-vapor triodes are also used. A mercury-vapor tube with a grid is called a thyratron. The construction of a thyratron is somewhat different from the construction of a vacuum-type triode tube. A sketch of a thyratron is shown in Fig. 34.

The control that the grid has over the flow of plate current in the thyratron is quite different from the control the grid has in the high-vacuum rectifier. In a thyratron, as long as the grid is maintained sufficiently negative to cut off the flow of plate current, there will be no electron flow from the cathode of the tube to the plate. Even with a high positive voltage on the plate of the tube, the grid, if it is negative enough, can block the flow of electrons to the plate. Electrons emitted by the cathode are simply driven back into the space charge and to the cathode.

Up to this point the action of the grid in a thyratron is similar to the action of the grid in a vacuum triode. However, if the grid voltage on the thyratron is reduced below the value required for cut-off, something entirely different happens.

A circuit using a thyratron is shown in Fig. 35A. If the grid is made negative and a positive voltage

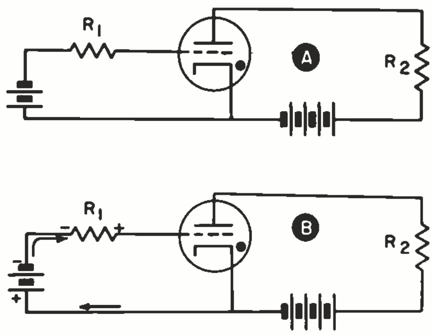


Fig. 35. Circuit using a thyratron.

is applied to the plate, the grid is able to cut off the flow of plate current. Once the grid voltage is reduced below cut-off, electrons begin flowing from the cathode of the tube to the plate. These electrons flowing through the tube will strike the gas molecules and knock electrons off them. These electrons will flow over to the plate of the tube. Meanwhile, the gas molecules that now have a positive charge on them will drift to the space charge to pick up electrons to neutralize the positive charge. In doing this, they remove the electrons in the space charge which in effect reduces the internal resistance of the tube, and permits a high current to flow.

If we try to cut off the flow of plate current by increasing the negative grid voltage to the cut off value again, the negative grid will attract positive ions which are gas molecules that have had electrons knocked off them. These positive ions will, in turn, attract electrons from the grid. These electrons are coming to the grid through R1, and as a result there will be a voltage drop across this resistor as shown in Fig. 35B. This voltage drop has a polarity opposite to that of the grid battery and will tend to neutralize the negative voltage applied in the grid circuit so the grid itself does not become very negative. Therefore, the grid is unable to gain control of the plate current, so the plate current continues to flow from the cathode to the plate of the tube even though the negative voltage applied to the grid circuit may be considerably greater than the applied voltage that would originally cut off the flow of plate current. If we increase the negative grid voltage to a very high value to try to cut off the flow of plate current, we simply attract more ions to the grid. These positive ions will attract more electrons from the grid, with the result that the number of electrons flowing from the grid to the ions may become quite high. This current could become so high the tube would be destroyed if it were not for the resistor R1 placed in the grid circuit which helps to keep the grid current down to a safe value.

In a thyratron, if the plate current is to be cut off by the grid voltage, the negative voltage needed to cut off the flow of plate current must be applied to the grid of the tube before the plate voltage is applied. Then the positive voltage can be applied to the plate of the tube, and the grid will prevent the electrons from reaching the plate of the tube. But once the grid voltage is reduced to a point called the "starting" point, where electrons can begin flowing from the cathode to the plate, the tube will fire, just as the mercury-vapor diode does, and the positively charged molecules will neutralize the space charge. This will permit a high current to flow through the tube and at the same time be attracted to the grid and neutralize any negative voltage placed on it to try to cut off the flow of plate current.

In a circuit using a thyratron, once the grid has lost control of the flow of plate current, the only way it can regain control is by removing the plate voltage from the tube. Once the plate voltage is removed, the grid can regain control; plate voltage can then be reapplied to the tube and no current will flow through it as long as the grid voltage is kept sufficiently negative to prevent any electrons from flowing from the cathode to the plate of the tube.

Thyratrons are extremely useful in industrial electronic applications. In some circuits, ac is applied to the plate of the tube instead of dc.

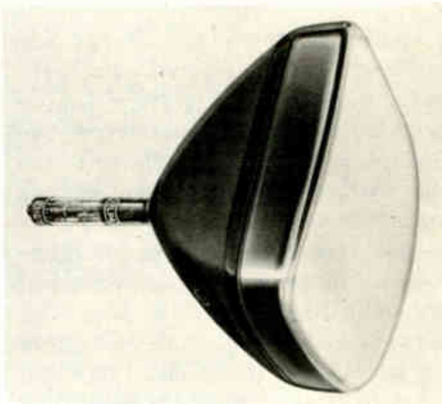
THE CATHODE-RAY TUBE

Another important tube found in many pieces of test equipment and also in TV receivers is the cathode ray tube. A photo of a cathode ray tube is shown in Fig. 36.

In the neck of a cathode ray tube is a device known as an electron gun. A typical electron gun is also shown in Fig. 36. The electron gun contains a heater, which heats a cathode; the cathode emits electrons; and these electrons are attracted by the high positive voltage on the anodes in the tube. Between the first anode and the cathode is the grid, as shown in the figure. The grid is used to control the number of electrons passing from the cathode towards the anodes.

The anodes in a cathode ray tube are arranged with a hole in the center of them so that instead of attracting electrons to themselves, they accelerate the electrons down the neck of the tube. The electrons pass right through the holes in the anodes and travel at a very high speed towards the face of the cathode ray tube. The face of the cathode ray tube is covered with a phosphorescent type of material. When this surface is struck by electrons, it glows and gives off light.

If the electrons accelerated down the electron gun were permitted to travel directly towards the face of the tube, they would all strike the face at approximately the same spot--somewhere near the center of the tube. However, the electron beam



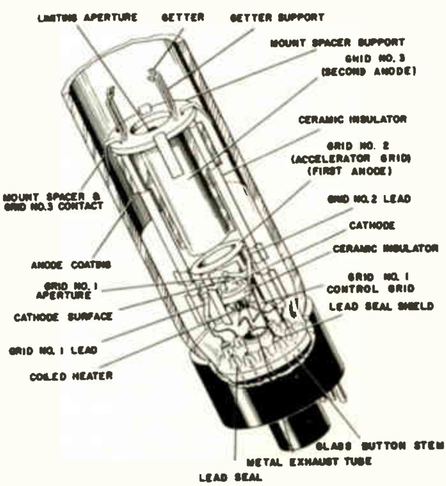


Fig. 36. A cathode ray tube is shown above; an electron gun of a crt below.

can be deflected so that it can be made to strike any point on the face of the tube.

The electron beam can be deflected by means of plates inside the neck of the cathode ray tube. If a positive voltage is put on one plate and a negative voltage on another, the positive plate will attract and bend the electron beam towards it and at the same time the negative plate will push the beam away from it. If two sets of parallel plates are installed in a tube, one in the horizontal plane and the other in the vertical plane, the horizontal plates can be used to move the electron beam up and down, and the plates installed in the vertical direction can be used to move the electron beam from side to side. Such a tube is called an electrostatic cathode ray tube.

The electron beam may also be deflected by magnetic fields produced by two pairs of coils. This is called electromagnetic deflection. The picture tubes used in TV receivers use this type of deflection.

Color picture tubes are similar to the tubes shown in Fig. 36 except that three electron guns are used inside of the tube. One gun is used for each of the three primary colors, red, blue and green. These guns are arranged to produce three electron beams that travel to the face of the picture tube and strike phosphors that will give off colored light. In other words, the blue gun strikes a phosphor dot that produces blue; the red gun strikes a phosphor dot that produces red, and the green gun strikes a phosphor dot that produces green. Hundreds of thousands of dots are placed on the face of the tube, and by sweeping the electron beams over the face of the tube it is possible to produce color pictures. Of course, the color tube is more complex than this, and there are other parts that we have omitted. We will study them later, but this will give you a general idea of what a color tube is like.

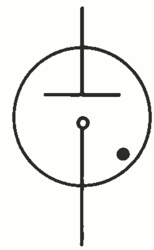


Fig. 37. Schematic symbol of a VR tube.

OTHER SPECIAL TUBE TYPES

There are many other special tube types found in electronic equipment.

Voltage Regulator Tubes.

One type that is quite common is the voltage regulator tube, often abbreviated the VR tube. The schematic symbol used to represent a VR tube is shown in Fig. 37.

Notice that the VR tube is a gas-filled tube. It is often used in a circuit like the one shown in Fig. 38. An important characteristic of this type of tube is that it maintains an almost constant voltage drop across it. If the voltage tends to increase, the tube will draw current, which will result in a greater voltage drop across the resistor so that the volt-

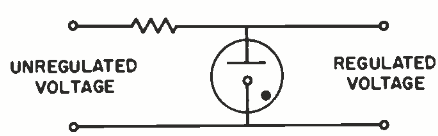


Fig. 38. A circuit using a VR tube.

age across the tube will remain constant. This type of tube is used to regulate the voltage in circuits where it is important for the voltage to be held as constant as possible.

Photo Tubes.

Another type of tube found in electronic equipment is the photo tube. This tube has a cathode and an anode, or plate. Instead of emitting electrons by thermionic emission, this tube is designed to emit electrons when light strikes its cathode. Electrons travel over to the plate of the tube and thus current flows through the tube. In Fig. 39 we have shown a picture of a photoelectric tube and the schematic symbol for it.

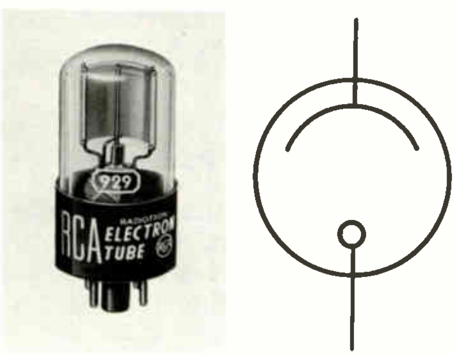


Fig. 39. A photoelectric tube and its schematic symbol.

SUMMARY

There are many special tube types in electronic equipment. Among these are the gas-filled tubes. Gas-filled diodes are frequently used in power supplies, particularly where the current drawn from the power supply goes through wide variations. Mercury-vapor rectifier tubes are found in many pieces of industrial electronic equipment and in the power supplies of most radio and TV transmitters.

A three-element mercury-vapor tube is called a thyratron tube. The most important characteristic of a thyratron is that the grid can prevent the flow of current from the cathode to the plate if it is made negative enough before the plate voltage is applied to the tube. However, once the current begins to flow from the cathode to the plate of the tube, putting a negative voltage on the grid of the tube will not cut off the flow of plate current because the negative voltage attracts the positive gas ions in the tube and these ions draw electrons from the grid. If the grid voltage is high enough, this will cause such a high grid current that the tube may be destroyed.

The cathode ray tube is found in TV and in many pieces of test equipment. It contains an electron gun which is used to shape the electrons into a beam and to accelerate them down the gun towards the face of the tube. Electrons striking the face of the tube cause the tube to glow and give off light. Cathode ray tubes are made in two types, those using electrostatic deflection and those using electromagnetic deflection.

Other special tubes are the photo tubes and the voltage regulator tubes. You will see more of these tube types later.

SELF-TEST QUESTIONS

(au) What do we mean when we say a mercury-vapor fires?

(av) What is the advantage of a mercury-vapor diode over a vacuum-tube diode as a rectifier?

(aw) What is a thyratron?

(ax) What is the main difference between a thyratron and a conventional triode tube?

(ay) What is the purpose of the anodes in the electron gun of a cathode-ray tube?

(az) What two types of deflection may be used with a cathode-ray tube?

LOOKING AHEAD

Now that you have studied the various tube types and have learned how they operate, the next thing is to see them actually used in typical circuits. In your next lesson you will study the most common circuits found in electronic equipment. These circuits are the basic circuits from which more complex circuits have been developed.

ANSWERS TO SELF-TEST QUESTIONS

(a) A cathode and a plate.

(b) Tube cathodes can be divided into directly heated and indirectly heated types.

(c) A filament.

(d) By means of a heater which is a coil of wire placed inside the cathode. The cathode is usually made in the form of a hollow tube.

(e) To provide an abundant supply of electrons at low operating temperatures.

(f) None. Other than to heat the cathode the heater serves no useful purpose and therefore it is often left off the schematic diagram in order to simplify the diagram.

(g) Approximately 8 volts. The first number or numbers preceding the first letter in a tube type indicates the approximate heater voltage required by the tube.

(h) The plate or anode.

(i) The plate receives a certain amount of heating from the cathode of the tube and the remainder is produced by electrons striking it.

(j) The plate may begin to emit electrons and the excessive heat may force gases out of the plate material into the space surrounding the plate of the tube.

(k) The getter is used to eliminate gases that are released inside the tube the first time it is heated after it has been sealed.

(l) A tube from which all of the gases have been removed.

(m) A tube into which a certain amount of gas has been deliberately introduced. An example of the type of gas frequently introduced into diode tubes is mercury vapor.

(n) It simply indicates that a few electrons are missing the plate and striking the glass. This often happens and does not mean that there is anything wrong with the tube.

(o) When we say the plate current-plate voltage is linear we mean that a given change in plate voltage will produce a constant change in plate current. In other words, if we increase the plate voltage from 70 volts to 80 volts it will cause a certain increase in plate current. If we increase the plate voltage another 10 volts it will cause a similar increase in plate current.

(p) The grid is much closer to the cathode than the plate and therefore the voltage applied to the grid of the tube has a greater effect on plate current than the same voltage applied to the plate will have.

(q) The positive voltage applied to the grid of the tube will cause the number of electrons flowing from the cathode to increase. Most of these electrons will flow through the grid structure to the plate because the plate will normally have a much higher positive voltage than the grid. By the time the electrons have reached the grid they will be travelling at such a high speed that they will pass right through the grid and flow over to the plate. However, some of the electrons will be attracted by the grid and will cause some current to flow in the grid circuit.

(r) As the negative grid voltage is increased, the plate current will decrease until eventually the grid voltage will become negative enough to prevent any electrons from reaching the plate. When this happens we say that the plate current is cut off.

(s) The amplification factor of a tube is the ratio of the change in plate voltage to the change in grid voltage required to produce the same change in plate current.

(t) Between 5 and about 100.

(u) The amplified signal will normally be the same as the input signal except it will be 180° out of phase. In other words, when the input signal reaches its maximum positive value, the amplified signal will reach its maximum negative value and when the input signal reaches its maximum negative value, the amplified signal will reach its maximum positive value. We say that the output signal is inverted or 180° out of phase with the input signal.

(v) The ac plate current is the changing plate current produced by the input signal voltage. It acts like an ac current superimposed on the dc plate current that flows through the tube when the input signal is zero.

(w) A grounded-cathode amplifier is an amplifier using a tube in which the cathode is at signal ground potential.

(x) An E_g-I_p curve is a grid voltage-plate current curve. It shows how the plate current varies with different values of grid voltage.

(y) The linear portion of the E_g-I_p characteristic curve is the straight portion of the characteristic curve. In Fig. 15, the linear portion is between approximately -6 volts and 0 volts.

(z) A grid-bias voltage is a voltage applied between the grid and the cathode of the tube to fix the operating grid voltage so that the tube will operate over the linear portion of the E_g-I_p characteristic curve.

(aa) When a tube is operated with too low a grid bias, the signal may drive the grid in a positive direction. When this happens, the grid will draw current with the result that the increase in plate current will not be linear. This results in a flattening of one half of the amplified output signal.

(ab) When the operating grid bias on a tube is too high, the input signal may drive the grid so far in a negative direction that the flow of plate current may be completely cut off. This will result in a flattening of the negative half of the plate current cycle which will produce amplitude distortion.

(ac) An Ep-Ip curve is a characteristic curve which shows the plate current for different values of plate voltage. A series of these curves are usually given, one curve for each value of grid voltage.

(ad) The ac plate resistance of a tube is the ratio of a change in plate voltage to a change in plate current that it produces. It is represented by the formula:

$$r_p = \frac{e_p}{i_p}$$

(ae) The dc plate resistance of a tube is the dc plate voltage measured between the plate and cathode of the tube divided by the dc plate current flowing through the circuit.

(af) The mutual conductance of a tube is equal to a change in plate current divided by the change in grid voltage required to produce the change in plate current. Mutual conductance is usually represented by g_m and the formula for mutual conductance is

$$g_m = \frac{i_p}{e_g}$$

(ag) The mutual conductance is measured in mhos. However, the mho is a rather large unit and thus we usually convert this to micro-mhos by multiplying it by 1,000,000.

(ah) $\mu = r_p \times g_m$.

(ai) An equivalent circuit is a circuit used to analyze the performance of an amplifier stage.

(aj) The generator voltage is $-\mu e_g$. The voltage is negative to indicate the fact that it is inverted by the stage. In other words, the output voltage is 180° out of phase with the input voltage.

(ak) The generator internal resistance is the plate resistance of the tube.

(al) The stage gain will be 45. To find the gain of the stage we use the formula

$$\text{stage gain} = \mu \times \frac{R_L}{R_L + r_p}$$

and substituting 50 for the amplification factor, 90,000 for the load resistance and 10,000 for the plate resistance we get

$$\text{stage gain} = 50 \times \frac{90{,}000}{90{,}000 + 10{,}000}$$

$$= 50 \times \frac{90{,}000}{100{,}000}$$

$$= 50 \times \frac{9}{10}$$

$$= 45$$

(am) The screen grid reduces the plate-to-grid capacitance and prevents oscillation due to feedback from the plate to the grid of the tube.

(an) The screen of a tetrode tube is operated with a positive potential applied to it. Usually the positive potential is about half the plate potential.

(ao) The screen is operated at signal ground potential. We accomplish this by connecting a suitable capacitor between the screen of the tube and ground. The capacitor offers a low reactance to ac signals on the screen so, insofar as the signals are concerned, the screen is essentially at ground potential.

(ap) The five elements in the pentode tube are the cathode, the grid, the screen grid, the suppressor grid and the plate.

(aq) The grid and the screen grid. The voltage applied to the grid and the voltage applied to the screen grid will control the flow of plate current in a pentode tube. The plate voltage on a pentode tube has very little effect on the plate current flowing in the tube.

(ar) In a beam-power tube the electrons are focused in two beams. Electrons knocked off the plate are moving at a slow speed and they are repelled by the high-speed electrons in the beam back to the plate of the tube.

(as) A Compactron is simply a tube where several types have been combined in just one envelope. The base of the Compactron tube has twelve pins so that it is possible to combine a number of complete tubes in the same envelope.

(at) A Nuvistor is a miniature tube shaped something like a thimble with a flat top. The tube is extremely small and therefore the capacity between the plate and grid is quite low, making the tube suitable for use in rf amplifiers.

(au) When we say a mercury vapor tube fires we mean that the gas inside the tube ionizes.

(av) The mercury vapor tube has a constant voltage drop regardless of the current flowing through the tube. The vacuum type rectifier does not have this desirable characteristic; the voltage drop across the tube will depend upon the current flowing through the tube.

(aw) A thyratron is essentially a triode tube that has been filled with a gas such as mercury vapor.

(ax) In a thyratron, the grid can keep the plate current cut off if a high negative voltage is applied to the grid before voltage is applied to the plate of the tube. Once the plate current begins to flow, the grid loses all control of the flow of plate current and normally cannot be used to reduce or cut off the plate current. In a vacuum-type triode tube, however, the grid always maintains control over the flow of plate current.

(ay) The anodes are used to accelerate and focus the electron beam. They are not designed primarily to attract electrons as is the plate in the conventional tube, but rather they are used to accelerate electrons in the form of a beam down the electron gun towards the phosphor on the face of the cathode-ray tube.

(az) Electrostatic deflection and electro-magnetic deflection.

Lesson Questions

Be sure to number your answer sheet B110.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time or you may run out of lessons before the new ones can reach you.

1. What type of emission is used in a vacuum tube with either a directly or an indirectly heated cathode?

2. Approximately what heater voltage would you expect a 12BE6 tube to require?

3. Why does secondary emission occur at the plates of most tubes?

4. If we find that in a certain triode tube a change in grid voltage of 2 volts produces the same change in plate current as a change in plate voltage of 100 volts, what is the amplification factor of the tube?

5. When the signal voltage swings the grid of the amplifier shown in Fig. 13 in a positive direction, in what direction does the plate voltage swing? Why?

6. Why do we apply grid bias to a tube?

7. Name the three important tube characteristics and give the formula for each.

8. Why can the stage gain of a triode amplifier never quite equal the amplification factor of the tube?

9. How does the suppressor in a pentode tube prevent the undesirable effects of secondary emission that occur in a tetrode?

10. How can you stop the flow of plate current in a thyratron once the tube has fired?

ENTHUSIASM

Starting work in a new field of endeavor just naturally arouses an intense and eager interest in what you are doing -- an enthusiasm which not only makes study and work a pleasure, but also betters your chances for success.

My students have enthusiasm for their Course of training because it is preparing them for a definite goal -- an independent business or a good job. They are continually finding, in the lessons studied, explanations for mysteries which they have encountered. Discovering immediate uses for fundamental facts keeps their enthusiasm high.

When enthusiasm is aroused in an ambitious man, he reacts like a thoroughbred race horse, giving all the speed and effort in him. If you try to arouse enthusiasm in a mule, however, all he will do is kick!

A real and lasting enthusiasm for electronics will make your study and work as pleasant as play, and will make your life much happier.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-210

ACHIEVEMENT THROUGH ELECTRONICS

HOW TUBES ARE USED

B111

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

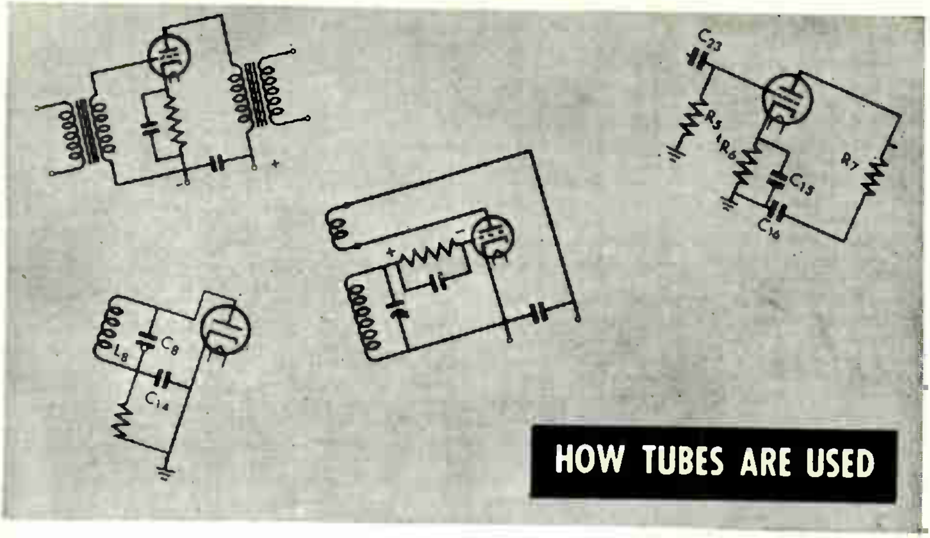
HOW TUBES ARE USED

B111

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

JDL3M1174 1974 EDITION Litho in U.S.A.



HOW TUBES ARE USED

In the preceding lesson you studied tube fundamentals and learned how a vacuum tube works. You learned that a tube is a unilateral device. By this we mean that it works in only one direction. Current will flow from the cathode to the plate of the tube, but it will not normally flow from the plate to the cathode.

You learned that in a three-element tube, a grid placed near the cathode can control the flow of electrons from the cathode to the plate. Because the grid is closer to the cathode than the plate is, it has a greater effect on the flow of plate current than the plate. Hence, a small signal voltage applied to the grid of a vacuum tube will cause the plate current to vary. This varying plate current can develop a voltage several times the original grid voltage across the plate load. This ability of the grid to control the flow of current from the cathode to the plate of the tube is what makes it possible for the tube to amplify a signal.

The purpose of this lesson is to increase your understanding of how tubes operate and to study a number of important basic tube circuits. You will study amplifiers similar to those used to amplify an audio or a video signal. You will study rf amplifiers, detectors and oscillators. All of these circuits will be found in modern electronic equipment. You will also study rectifiers. Although tube-type rectifiers are no longer used in new receiving-type equipment, they are still used in transmitters, in industrial applications and in many older radio and TV receivers.

These are the basic circuits that you are most likely to encounter. Later on, you will study many other circuits.

After we have looked at the basic circuits separately, we will see how these different circuits are used together to form a complete radio receiver. The basic principles used in radio receivers are similar to those used in both black and white and color television reception.

In the last section of the lesson we will take up some variations in amplifier circuits. In the simple amplifier you studied in the preceding lesson, the load was placed in the plate circuit of the tube and the cathode was operated at signal-ground potential. This type of amplifier is called a grounded-cathode amplifier. However, there are different circuits in which one of the other elements is operated at ground potential. In one type of amplifier the plate is grounded, and the load is placed in the cathode circuit. This type of amplifier is called a grounded-plate amplifier or more frequently a cathode-follower. In still another type of amplifier the grid is operated at ground potential, and the input signal is applied between the cathode and ground. The load is placed in the plate circuit. This type of amplifier is called a grounded-grid amplifier. In this lesson you will study all three types. While most of the amplifiers you will find are grounded-cathode amplifiers, grounded-grid amplifiers are used both in transmitting and receiving equipment. Also, cathode-followers are quite widely used in industrial applications, and they have been quite widely used in color TV receivers. The material you will study in this lesson is primarily an introduction to the different types of amplifiers. You will study them all in greater detail later.

Types of Amplifiers

Amplifiers can be divided into classes according to the amount of bias applied to the tube. You will remember that the bias is a negative voltage applied between the grid and cathode of a tube. If the bias applied to a tube is midway between zero bias and cut-off bias, the amplifier is called a Class A amplifier. A Class A amplifier operates on the linear or straight portion of the characteristic curve. Most of the amplifier stages found in radio and TV receivers are Class A amplifiers.

In some, the dc operating bias applied to a tube is equal to cut-off bias. This simply means that the negative voltage applied to the grid of the tube reduces the plate current to zero or almost to zero. This type of amplifier is called a Class B amplifier. It is found in medium and high power audio and video amplifiers and in radio and TV transmitters.

In still another type of amplifier the grid voltage applied to the tube is several times the negative voltage required to cut off the flow of plate current. This type of amplifier is called a Class C amplifier. Class C amplifiers are used as radio-frequency power amplifiers and as oscillators. You will see examples of all three classes of amplifiers in this lesson.

Amplifiers can also be divided into two general types, voltage amplifiers and power amplifiers. In a voltage amplifier we are interested in amplifying the signal voltage. A weak signal voltage is applied to the input circuit of the stage, and we are interested in getting as much amplification as possible. In other words we want as high an amplified voltage in the output as possible.

Usually, in order to get a high output voltage, the load used is a fairly high impedance or resistance.

In a power amplifier we are not particularly concerned about the amplitude of the voltage in the output. Instead, we are interested in getting as large a current variation from the tube as possible. Usually the load impedance is much lower in a power amplifier than it is in a voltage amplifier. However, in many cases it is difficult to tell a voltage amplifier from a power amplifier simply by looking at the circuit unless you know something about the value of the components used in the circuit and the conditions under which the tube is operated.

Voltage amplifiers are Class A amplifiers; power amplifiers may be Class A, B, or C.

VOLTAGE AMPLIFIERS

A simplified schematic of a typical voltage amplifier is shown in Fig. 1. Here the signal source is applied in the grid circuit between the grid and the cathode of the tube in series with the bias voltage. In the plate circuit of the tube we have the load connected in series with a B supply battery. The load and battery are connected between the plate and cathode of the tube. With a high-impedance load, small current variations flowing between the cathode and plate of the tube and through the load will result in comparatively large voltage variations across the load.

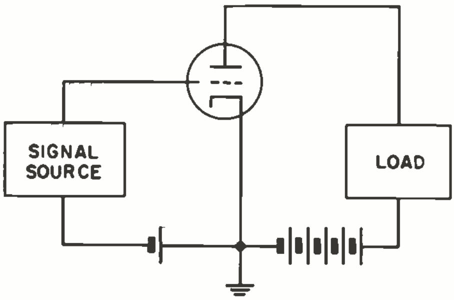


Fig. 1. A grounded-cathode amplifier.

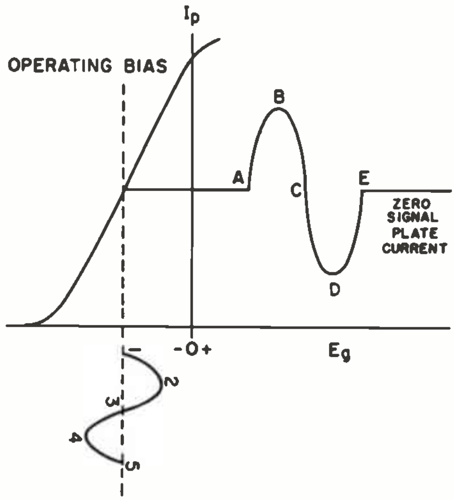


Fig. 2. Operating conditions of a typical voltage amplifier.

The graph in Fig. 2 shows how a voltage amplifier operates. Sufficient bias is applied to the tube to place the operating bias on the grid approximately midway between 0 and cut-off voltage. The idea is to bias the tube so that the tube will be operating on the linear portion of its characteristic curve. If we have too little bias on the tube, the grid may be driven positive, which will cause the grid to draw current; also the variations in plate current will not be linear. Similarly if too much bias is put on the tube, the grid will be driven beyond the cut-off point and there will be no plate current flow at all during part of the cycle of the input signal.

With the tube operated as shown in Fig. 2, the plate current variations follow the grid voltage. If the input signal swings in a positive direction it subtracts from the grid bias, making the bias less negative. This is what happens as the input moves from point 1 to point 2. As the input signal swings between these two points, the plate current moves from point A to point B.

During the next quarter of the input cycle, the grid voltage is swinging in a negative direction because the input signal is decreasing and dropping to zero as it moves from point 2 to point 3. At this point the signal has dropped to 0, and the voltage applied to the grid at this instant is the operating bias. During this quarter of the cycle the plate current moves from point B to point C.

During the next quarter cycle the input signal is swinging negative as it moves from point 3 to point 4. The signal now adds to the grid bias, making the voltage more negative. This results in the plate current dropping still further from point C to point D. When the input signal voltage reaches point 4 and starts to swing back to 0 again towards point 5, the grid becomes less negative, with the result that the plate current begins to increase and move from point D to point E.

There are several important things to be noted from the curve shown in Fig. 2. First, notice that the operating bias is approximately midway between 0 grid voltage and the grid voltage required for plate current cut-off. This places the operation of the tube on the linear or straight portion of the characteristic curve. As we have mentioned, an amplifier operated in this way, that is biased midway between zero voltage and cut-off voltage, is called a Class A amplifier.

Also notice that the input signal is small enough so that it neither drives the grid into the positive region nor does it drive it beyond cut-off. If the grid is driven positive or beyond cut-off, distortion will result, and the output signal will not be a faithful reproduction of the input signal.

Let us consider what happens to the plate current flowing in the tube. The line marked plate current represents the plate current that will flow when there is no input signal. Notice what happens when an input signal is applied. During one half cycle the plate current flow increases; during the other half cycle it decreases. The increase during one half cycle is equal and opposite to the decrease during the next half cycle. Therefore, if we consider the average plate current, it does not change. In other words, although the plate current does increase during one half cycle, it decreases by an equal amount during the next half cycle; the average plate current flowing remains the same as it was when no signal was applied to the input. Thus, if a dc milliammeter is placed in the plate circuit of a Class A amplifier, there will be no change in the reading when a signal is applied to the input because the meter will indicate the average dc flowing, and this does not change. If there is a change in the plate current, either the bias on the tube is incorrect or else the signal applied to the input is too strong and is driving the tube onto the non-linear portion of the characteristic curve.

The voltage amplifiers found in modern electronic equipment are Class A amplifiers. Thus, in a voltage amplifier you can expect to find that the output signal is a faithful reproduction of the input signal.

Furthermore you can expect to find the bias midway between zero bias and cut-off bias and also you can expect the average plate current flowing in the stage not to change when a signal voltage is applied to the input.

POWER AMPLIFIERS

Class A Amplifiers.

Class A amplifiers are used as power amplifiers but usually only when the amount of power required is comparatively low. The reason for this is that the Class A amplifier has relatively poor efficiency. The efficiency is the ratio of the power output to the power input.

It is easy to see why the efficiency of a Class A amplifier is poor. First, let's consider the operating curves shown in Fig. 3. Let's assume the tube is operating with a voltage of 200 volts applied to it and a plate current of 100 milliamperes. Thus the power input to the tube is 200 × .1 = 20 watts. This is the dc power input; it represents the power being taken from the power supply and fed to the tube.

Now let us consider how much signal power we can get out of this tube. If the current flowing in the tube is 100 milliamperes as shown in Fig. 3A, the maximum current change we can get in a half cycle is 100 milliamperes. In other words, if the grid voltage is swung all the way to the cut-off point by the input signal, then the plate current would drop to zero. This means that the peak current change is 100 mils, or .1 amp. The rms or effective current change, in amperes, will be only .707 of this value, or .1 × .707.

Similarly, the plate voltage applied to the tube is 200 volts as shown in Fig. 3B. The maximum change that could be made in plate voltage is 200 volts. To do this, the plate voltage would have to drop all the way to zero as the current through the load changed. Again this is the peak change, and the rms or effective change, in volts, is 200 × .707.

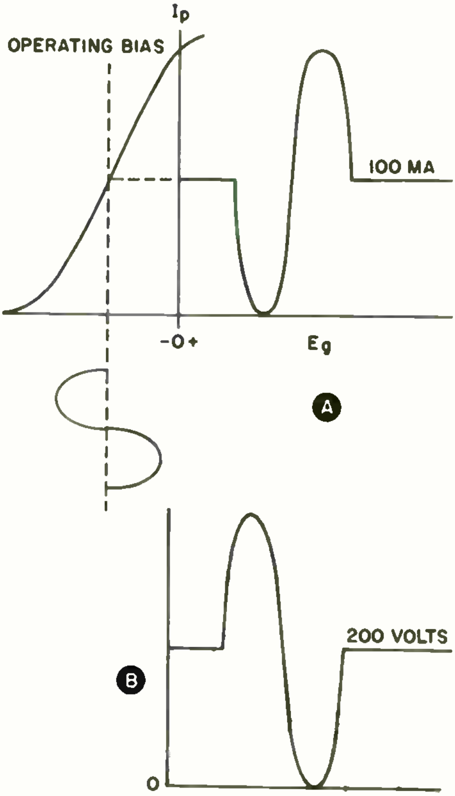


Fig. 3. Operating conditions for Class A operation.

The maximum ac power output is equal to the effective ac voltage times the effective ac current. Therefore, the power output is equal to:

$$200 \times .707 \times .1 \times .707$$

If we multiply .707 × .707 we have .5, so we can express the formula as: 200 × .1 × .5. Now you will recognize the 200 × .1 as the input power, and we know that .5 is equal to 50%. Thus, the maximum power output that can be obtained from a Class A power amplifier is equal to 50% of the input power. Therefore, the best efficiency that can be obtained from a Class A amplifier is 50%. In fact, it is usually impossible to obtain this high an efficiency under actual operating conditions, because the grid would have to be driven all the way to cut-off on one half cycle, and all the way to zero on the other half cycle. Under these conditions, considerable distortion would result. Actually, the plate current would not be a sine wave as shown in Fig. 3, but would be flattened somewhat on both top and bottom. Efficiency of somewhere around 30% to 35% is usually about all that can be obtained from a Class A amplifier if reasonable linearity is to be maintained.

Class B Amplifiers.

Better efficiency can be obtained from a power amplifier if the bias on the tube is increased to approximately cut-off bias and a signal large enough to drive the grid positive is used. Under these conditions the operating curves for the tube will look like those shown in Fig. 4. When there is no signal applied to the input of the stage, the plate current flowing through the tube will be low. When the input signal swings in a positive direction, plate current flows in the form of a large pulse. When the signal swings in a negative direction, it soon drives the grid beyond the cut-off point and there will be no plate current flow in the tube. Thus you can see that the only time the tube is called on to furnish a large amount of power

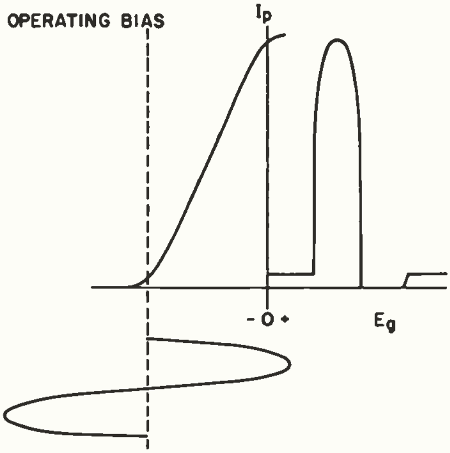


Fig. 4. Operating conditions for Class B operation.

is when the grid is driven in a positive direction. For approximately half of each cycle there is no current flow through the tube at all.

It is immediately apparent that with this type of operation only one-half of the input signal is being reproduced. In an audio amplifier this would result in a great deal of distortion. However, this problem can be overcome by the use of two tubes, one to reproduce each half of the audio signal. By recombining the output from these two tubes, both halves of the audio signal can be reproduced. We will see examples of these circuits later in this lesson.

It is easy to see that this type of amplifier is more efficient than a Class A amplifier, because the current flowing through the tube when there is no signal present is comparatively small. Almost all of the current capabilities of the tube are reserved for the reproduction of the output signal. This type of operation, with the tube biased to cut-off, is called Class B operation. It is used quite extensively where large amounts of audio or video power must be developed, and it is also used to develop rf power in certain types of rf equipment.

Class C Amplifiers.

A still more efficient power amplifier than a Class B amplifier is the Class C amplifier. The curve shown in Fig. 5 shows how a tube is operated as a Class C amplifier. Notice that the operating bias is greater than the bias required to cut-off the flow of current through the tube. The Class C amplifier tube is generally operated with a bias somewhere between 2 and 4 times cut-off bias.

The input signal required for a Class C amplifier is considerably higher than that required for a Class A or Class B amplifier. The input signal drives the grid well into the positive region so that the grid draws substantial current. Plate current flows only in a series of pulses and these pulses actually flow for considerably less than half a cycle.

Class C amplifiers are not used for audio work, but they can be used as radio-frequency power amplifiers. They are used in conjunction with resonant circuits. If a parallel resonant circuit is placed in the output of a Class C amplifier, the pulse

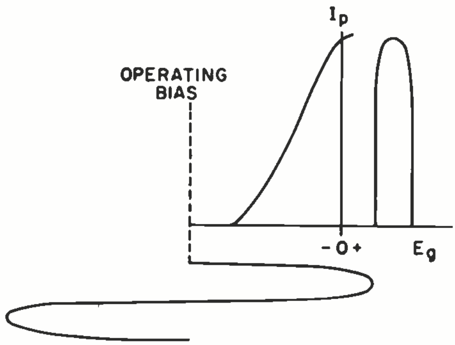


Fig. 5. Operating conditions for Class C operation.

from the Class C amplifier will shock-excite the parallel resonant tank circuit into oscillation. If the tank circuit receives a pulse once each cycle, this pulse fed into the tank circuit is able to make up any losses in the tank circuit. Meanwhile, the current flows back and forth in the tank circuit so that the voltage appearing across it is actually a sine wave, even though it is being supplied energy only in the form of pulses. The tank circuit in a Class C amplifier has sort of a flywheel effect, and once oscillations are set up in it they can be maintained by pulsing it once each cycle.

As a matter of fact, the tank circuit can often be designed so that it does not need a pulse once each cycle but can maintain oscillations by receiving a pulse every other, or perhaps every third cycle. By designing the tank circuit in this way, it is possible to double or triple the frequency of the signal.

In addition to Class A, Class B, and Class C amplifiers there are also amplifiers known as Class AB amplifiers. As the name suggests, these are simply amplifiers operated midway between Class A conditions and Class B conditions. The operating bias applied to the tube is a little higher than that needed for Class A operation, but not as high as that needed for Class B operation. Beam power tubes are often operated under Class AB conditions.

SUMMARY

The important points to remember from this section are that there are three classes of operation for vacuum tubes. A tube that is operated with a bias midway between zero bias and cut-off bias is a Class

A amplifier. The output of this type of amplifier should be an exact duplicate of the input.

The Class B amplifier is operated at approximately cut-off bias. The zero-signal plate current flowing in this type of stage is low, and only the positive half of each cycle is reproduced. A Class C amplifier is an amplifier operated with bias several times cut-off bias. Plate current flows in this type of stage for less than half of each cycle.

Voltage amplifiers are Class A amplifiers. Power amplifiers may be Class A, Class B, or Class C. A Class A power amplifier has relatively poor efficiency. The efficiency of a Class B amplifier is better than that of a Class A amplifier, and the efficiency of a Class C amplifier is still better than the efficiency of a Class B amplifier. Class AB amplifiers are amplifiers operated under conditions between Class A and Class B.

SELF-TEST QUESTIONS

(a) What is a Class A amplifier?
(b) What is a Class B amplifier?
(c) What is a Class C amplifier?
(d) Into what two general types can amplifiers be divided?
(e) What is the purpose of a voltage amplifier?
(f) Into which class does the voltage amplifier fall?
(g) What is the disadvantage of a Class A power amplifier?
(h) What is the maximum possible efficiency that can be obtained from a Class A power amplifier? What is the practical efficiency of a Class A amplifier?
(i) Can a Class B power amplifier be used in audio power amplification?
(j) What type of signal is the Class C power amplifier used to amplify?
(k) What is a Class AB amplifier?

Typical Amplifiers

Look at Fig. 1 again, where we have shown a simplified diagram of an amplifier. As you can see from this diagram, the basic parts of an amplifier are the tube, the load, and batteries or some other power source to supply the power needed to operate the tube. Now, let us study some practical amplifiers to see how the various electronic components you have already studied are used in conjunction with tubes in order to amplify signals.

AUDIO AMPLIFIERS

An audio amplifier is an amplifier designed to amplify signal frequencies within the range normally heard by our ears. Some amplifiers are capable of doing this job better than others. The audio amplifier found in the average radio or TV receiver is not capable of reproducing all of the frequencies that our ears can hear. However, an amplifier designed for use in a piece of high-fidelity equipment has a much better frequency response and can amplify a wider range of frequencies. Amplifiers found in most radio and TV receivers can amplify frequencies from about 50 or 60 cycles per second up to 8000 or 9000 cycles per seoond. Amplifiers designed for use in high-fidelity equipment can amplify frequencies from about 10 cycles per second up to at least 15,000 cycles per second and sometimes as high as 100,000 cycles per second.

Audio amplifiers may be either voltage amplifiers or power amplifiers. Voltage amplifiers are used to build up the strength of the weak audio signal until it is strong enough to drive a power amplifier. A power amplifier is then used to supply the power to drive the speaker.

Voltage Amplifiers.

A typical voltage amplifier is shown in Fig. 6. This amplifier is called a resistance-capacitance coupled amplifier because resistors and capacitors are used to couple the signal to the amplifier and to the output or to the following stage.

In the circuit shown in Fig. 6, capacitor C1 is used to couple the signal source to the grid of the tube. C1 will block any dc in the input circuit and keep it away from the grid of the tube. At the same time, if C1 is large enough, it will offer a low reactance to the flow of an ac signal through it and act, as far as the ac signal is concerned, as though it were not there at all. The input signal is therefore applied between the grid of the tube and ground.

Resistor R1 is called a grid leak. Some of the electrons travelling from the cathode to the plate of the

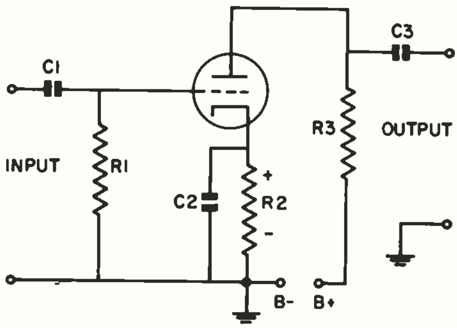


Fig.6. A typical resistance-capacitance coupled amplifier.

tube will accidentally strike the grid. If there is no way for these electrons to get off the grid of the tube, they will be trapped on the grid and eventually build up a high negative charge on the grid. This negative charge will reduce the flow of current from the cathode to the plate of the tube. As a matter of fact, in some tubes this charge may become so high it can actually cut off the flow of electrons from the cathode to the plate of the tube.

As long as R1 is in the circuit, electrons striking the grid can flow through R1 back to ground and then through R2 back to the cathode of the tube. Of course, when electrons flow through a resistor they build up a voltage across the resistor. Electrons flowing from the grid of the tube back to ground will develop a dc voltage across R1 having a polarity such that the grid end of the resistor is negative. Normally in an amplifier of this type, however, the number of electrons flowing through the grid resistor is not large enough to develop any appreciable voltage across the grid resistor, R1, even though the value of the resistor may be quite large. Resistors of 100,000 ohms to 500,000 ohms are frequently used as grid resistors. Usually it will be impossible to detect any voltage across these resistors, even with quite sensitive measuring equipment.

Capacitor C1 and resistor R1 actually form a voltage-divider network that divides the ac signal voltage. If R1 is made large in comparison to the reactance of C1, most of the signal voltage will appear across R1. This means that most of the signal voltage will be applied between the grid of the tube and ground, which is essentially where we want it. On the other hand, if R1 is small compared to the reactance of C1, then C1 and R1 will divide the signal voltage so that only a small part of it will appear across R1; the remainder will be lost across C1. This situation is to be avoided if maximum output is to be obtained from the amplifier. For this reason, designers make R1 as large as practical.

While we are discussing the combination of C1 and R1, remember that the reactance of a capacitor depends upon the frequency of the signal voltage. As the frequency decreases, the reactance increases. Therefore, at low frequencies the reactance of C1 may become large enough to appreciably reduce the voltage across R1. When this happens, the signal amplification will fall off; the amplifier will not amplify low-frequency signals as well as it does higher-frequency signals. Making R1 large tends to extend the low-frequency gain of an amplifier.

On the other hand if R1 is made too large, then the electrons accidentally striking the grid of the tube will develop an appreciable dc voltage across R1 when they flow through it. In an amplifier like the one shown in Fig. 6, this voltage is undesirable. Therefore the value of R1 must be a compromise. It is made as high as possible, so that the signal voltage will be high, without making it so high that a troublesome dc voltage will be developed across it.

The dc voltage supply used to supply the plate voltage to the tube is represented by the terminals B- and B+. This could be a B battery or it could be the terminals of a power supply. A power supply is a unit that converts ac from the power line to dc for use in applications such as this. You will study power supplies shortly.

Resistor R2 is put in the cathode circuit of the tube to eliminate the need for a C battery to supply a grid bias voltage. Let's consider what happens with this resistor in the circuit. Electrons flowing through the tube are emitted by the cathode, attracted by the plate, then flow through R3, then through the B supply to ground, and finally through R2 back to the cathode. The minute power is applied, this action is instantaneous, and electrons start flowing in all parts of the circuit. The electrons flowing through R2 will develop a voltage across this resistor with the polarity indicated on the diagram. This makes the cathode slightly positive with respect to ground. If the tube is designed to operate with a grid voltage of -3 volts, R2 is selected so that the electrons flowing through it will develop a voltage of 3 volts across the resistor. This will make the cathode 3 volts positive with respect to ground.

To see how this voltage biases the tube and eliminates the need for a C battery, let us consider the potential of the grid with respect to ground. The number of electrons flowing through R1 is so small that little or no voltage is developed across it. Therefore, the grid is normally at ground potential. This means that the cathode is positive with respect to the grid. If the cathode is positive with respect to the grid, then the grid is negative with respect to the cathode.

R2 is often called a cathode bias resistor. By using this resistor in the cathode circuit, we can eliminate the need for a C battery. Batteries deteriorate and have to be replaced periodically, whereas the resistor will last almost indefinitely providing it is not overloaded.

Capacitor C2 is connected across R2 in order to stabilize the voltage across R2. Without C2, as the input signal caused the tube current to vary, the current through R2 would vary. This would result in varying voltage or varying bias across the resistor. To eliminate this effect, we connect capacitor C2 across the resistor. C2 must be large enough to maintain the voltage across R2 constant. It does this by charging when the current through R2 increases and the voltage tends to rise, and discharging through R2 when the voltage tends to fall. The capacitor actually acts as a low reactance path for the ac signal through it. The grid of the tube, in causing the plate current to vary, is actually producing an ac signal superimposed on the dc in the plate-cathode circuit. This ac signal component flows through C2; the dc component flows through R2.

Resistor R3 is the plate load resistor. The value of this resistor is usually quite large. The larger the resistor, the closer the gain of the stage will approach the amplification factor of the tube. You will remember that the tube acts like a generator, and this generator has an internal resistance, the ac plate resistance of the tube. Resistance R3 is, in effect, connected in series with the plate resistance of the tube, and the voltage developed by the tube is divided between the plate resistance and the plate load. By making the plate load resistor as large as possible, we will get as much of the amplified signal voltage across this resistor as possible.

However, there is a limit to how large we can make this resistor.

Since the plate current flows through the resistor, there will be a voltage drop across the resistor. This voltage drop subtracts from the supply voltage so that the net voltage available to operate the tube is equal to the supply voltage minus the voltage drop across the plate load resistor. If we make the plate load resistor too large, there will be a very high voltage drop across it with the result that there is very little voltage left to operate the tube. Again, the selection of the size of resistor to be used is a compromise. A value is chosen that will give a reasonably high gain without an excessive voltage drop. If the voltage drop across R3 is excessive, then the power supply voltage must be very high in order to get the voltage we need on the plate of the tube. This could require a costly power supply, so it is often more economical to use two stages to get the gain we need than to try to get it from one stage by using an excessively large plate-load resistor.

A schematic diagram of another voltage amplifier is shown in Fig. 7. This amplifier is called a transformer-coupled amplifier. Notice that in some respects the circuit is similar to the circuit shown in Fig. 6. In the transformer-coupled amplifier, transformer T1 is taking the place of C1 and R1 in the

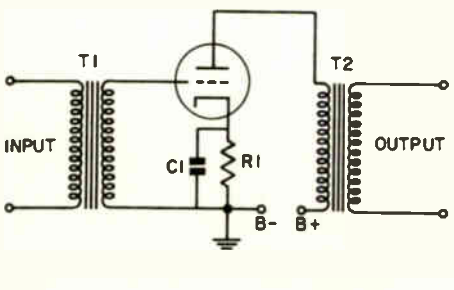


Fig. 7. A transformer-coupled amplifier.

resistance-coupled stage. Similarly, transformer T2 is taking the place of load resistor R3 and blocking capacitor C3 in the output. Both transformers can be step-up transformers, so that there will be voltage amplification in the transformers themselves as well as in the tube. Bias for the stage is still obtained by placing a resistor in the cathode circuit of the tube. R1 in Fig. 7 is the cathode-bias resistor, and it is bypassed by the capacitor C1.

Transformer-coupled amplifiers of this type are not used in modern electronic equipment. The transformers are more expensive than the resistor-capacitor combination, and modern tubes have such high gain that it is not necessary to rely on the step-up transformer to get a reasonable gain in the stage. However, you may be called on to service an older piece of equipment that might employ a transformer-coupled stage, so you should be aware that this type of coupling exists.

Power Amplifiers.

Fig. 8A shows a power amplifier using resistance-capacitance coupling in the input circuit. The input circuit is essentially the same as the circuit used in Fig. 6.

The output transformer is a step-down transformer. This transformer is primarily an impedance-matching device. It is used to match the low-impedance speaker, which would be connected across the output terminals, to the plate circuit of the tube. Remember that the tube works like a generator, and maximum power transfer will be obtained when the load matches the generator. The transformer matches the load impedance to the generator or tube impedance.

Fig. 8B shows a transformer-

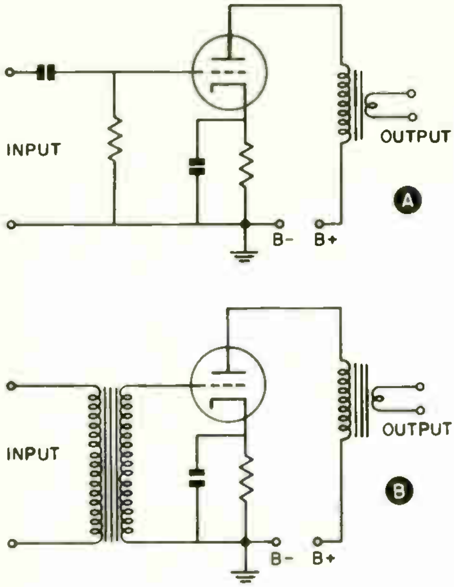


Fig. 8. Two single-ended power amplifier stages.

coupled power amplifier. The input transformer is a step-up transformer, and the output transformer will be a step-down transformer to serve as an impedance-matching device. The circuit shown in Fig. 8B is obsolete; you will not run into a circuit of this type except in old equipment.

Both of the amplifiers shown in Fig. 8 are Class A power amplifiers. They are also called single-ended stages because each circuit uses a single tube. In some audio amplifiers, in large radio receivers, and in some TV sets you will run into a double-ended power output stage such as the one shown in Fig. 9. This circuit is called a push-pull amplifier. It can be operated as either a Class A, Class AB, or a Class B power amplifier. If the stage is operated as a Class A power amplifier, the stage preceding it can be a voltage amplifier and Tl may be a step-up transformer. On the other hand, if the stage is operated as a Class B amplifier the stage preceding it must be a power amplifier because power must be supplied to the grid circuit, and T1 must be a step-down transformer.

If the stage is a Class AB_1 power amplifier, the preceding stage can be a voltage amplifier since the grid does not draw current and no power is consumed in the grid circuit. On the other hand, if the stage is a Class AB_2 power amplifier, the grid does draw grid current and power must be supplied to the grid circuit. The preceding stage, therefore, must be a power amplifier.

This stage is called a push-pull amplifier because it acts as though one tube is pushing electrons through the primary of T2 while the other is pulling electrons in the opposite direction. Briefly, the operation of

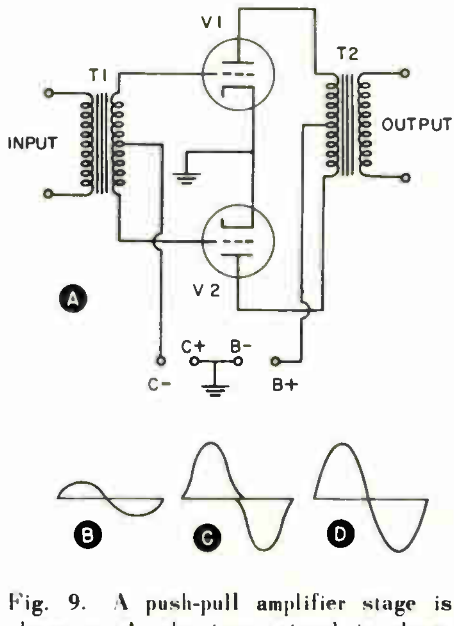


Fig. 9. A push-pull amplifier stage is shown at A, the input signal is shown at B, the output pulses from the two tubes at C, and the combined output at D.

the stage is as follows: the secondary of transformer T1 is tapped, and the center tap is at signal ground potential. When the end of the secondary of T1 that is connected to V1 is swinging positive with respect to ground, the other end will be swinging negative with respect to ground. The positive voltage applied to tube V1 will cause its plate current to increase, while the negative voltage applied to the grid of V2 will cause its plate current to decrease. Thus the current in one half of T2 increases while the current in the other half of T2 decreases. During the next half cycle, when the end of the secondary of T1 that is connected to V2 is positive with respect to ground, the other end will be negative. At this time the plate current of V2 will increase while the plate current of V1 decreases.

When this type of stage is used as a Class B amplifier, the tube that is driven positive conducts current heavily while the other tube does not conduct current at all. During the next half cycle the second tube carries the whole load while the other tube rests. If the input signal is a sine wave like that shown in Fig. 9B, the plate currents for the two tubes look like Fig. 9C and combine to produce a signal like Fig. 9D in the secondary of transformer T2. This explains how two tubes can be used in a Class B amplifier to amplify an audio signal when each tube conducts during only half of each cycle. One tube reproduces one half of the cycle; the other tube reproduces the other half of the cycle. The two signals are combined in transformer T2 to give an output signal that is an amplified reproduction of the input signal.

Push-pull amplifiers are used in some radio and TV receivers. They are found in high-fidelity equipment and in many radio and television transmitters. Push-pull amplifiers are used wherever it is necessary to develop a large amount of audio or video power. Operating these tubes as Class B amplifiers gives much better efficiency than operating them as Class A amplifiers. As a matter of fact, the same amount of audio power can usually be developed more economically by using two small tubes operated as Class B amplifiers than by using one large tube operated as a Class A amplifier.

RADIO-FREQUENCY AMPLIFIERS

Radio-frequency amplifiers like audio amplifiers can be divided into two types, voltage amplifiers and power amplifiers. The radio-frequency amplifiers found in receiving equipment are voltage amplifiers, whereas those found in transmitting equipment are power amplifiers. In receiving equipment, we are interested in taking the weak radio-frequency signal picked up by the antenna and amplifying it in order to extract whatever intelligence it may carry. In transmitting equipment, we are interested in developing power to feed to the antenna in order to radiate a strong signal.

Voltage Amplifiers.

A radio-frequency voltage amplifier is shown in Fig. 10. Notice that in many respects it is similar to the single-ended transformer-coupled audio amplifier. In the rf amplifier, we have used a pentode. Triodes are not as suitable as pentodes in rf amplifiers in most cases. Notice that both the input and output circuits are tuned. These circuits are

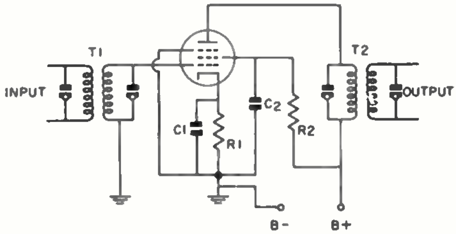


Fig. 10. A radio-frequency voltage amplifier.

adjusted to resonance at the frequency of the rf signal.

As in the audio amplifiers we studied, operating bias for the stage is obtained by inserting a resistor in the cathode circuit. This is the resistor marked R1 on the diagram; it is bypassed by capacitor C1. The purpose of capacitor C2 is to ground the screen of the tube insofar as signal voltages are concerned. Capacitor C2 is selected so that its reactance is low at the operating frequency. Thus, insofar as the signal is concerned, the screen is in effect operating at ground potential. This isolates the plate from the grid of the tube so that there is not enough energy fed from the plate of the tube back to the grid to cause the tube to go into oscillation. Resistor R2 is called the screen dropping resistor. Its purpose is to drop the B supply voltage to a suitable value for the screen. In many voltage amplifiers of this type, the screen voltage is somewhat less than the plate voltage. Plate voltage is applied to the tube through the parallel resonant circuit installed in the plate circuit.

With modern pentode tubes, a comparatively high voltage gain can be obtained in a stage of this type. It is easy to get a gain on the order of 100.

Power Amplifiers.

Radio-frequency voltage amplifiers are biased to operate at the mid-point of the characteristic curve. In other words, they are Class A amplifiers. However, rf power amplifiers are usually operated either in Class B or in Class C, although some rf power amplifiers operate in Class AB. A schematic diagram of a Class C rf power amplifier is shown in Fig. 11. Notice that this circuit differs somewhat from the voltage amplifier.

In this circuit, bias is obtained by means of a resistor in the grid instead of the cathode circuit. In a Class C amplifier a high value of bias is used. The signal applied to a Class C stage must be sufficient to drive the grid positive. When the grid is driven positive, electrons will leave the cathode and strike the grid to charge C2 with the polarity shown. During the time when the input signal is not positive, C2 will discharge through RFC and R1, making the grid negative with respect to ground. By selecting the proper value of R1, the correct bias can be developed across this resistor.

Since the tube in a Class C amplifier is normally operated at bias voltages several times cut-off, plate current does not flow through the tube except when the input signal

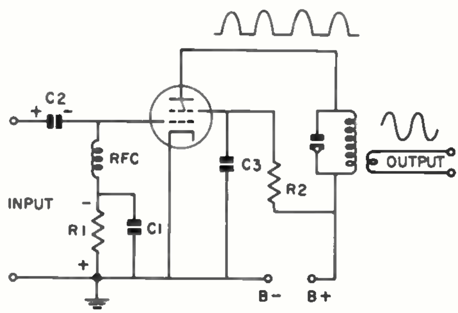


Fig. 11. Schematic of a Class C rf power amplifier.

swings positive and drives the grid into the region where plate current can flow. Current then flows from the cathode to the plate of the tube in the form of a series of pulses as indicated. These pulses shock-excite the parallel resonant tank circuit, consisting of the primary of the output transformer and the capacitor across it, so that current flows back and forth between the coil and capacitor, producing the sine wave output shown.

Class C rf power amplifiers have very high efficiency--in the range of 75%. This means that for each watt that is fed into the amplifier plate circuit, about 3/4-watt of rf power can be developed. The efficiency of this type of amplifier is somewhat better than that of a Class B amplifier and is much better than that of a Class A amplifier.

The stage shown in Fig. 11 is a single-ended stage. Push-pull Class C rf amplifiers can also be used. The power output of this type of stage is approximately double what could be obtained from a single-ended stage using the same tube type.

EFFICIENCY OF AMPLIFIERS

The amount of power you can get from a tube depends on which class of amplifier stage it is used in. As you have learned, a Class A stage has a practical efficiency of about 30 per cent. This means that 70 per cent of the power fed to the stage is lost. Most of this power is wasted at the plate of the tube. We say it is "dissipated" by the tube.

All power tubes have a rating known as the plate dissipation. This rating tells how much power can be dissipated at the plate of the tube without overheating the tube. Consider a power tube with a plate dissipation rating of 10 watts. If we use this tube in a power amplifier, we cannot let the plate dissipate more than 10 watts. This means that if the amplifier is a Class A amplifier with an efficiency of 30 per cent, 70 per cent of the input must not exceed 10 watts. The total power input to the stage must not exceed:

$$10 \times \frac{100}{70} = 14.28 \text{ watts}$$

This means the total power input to the stage should be about 14 watts, and the power output will be about 4 watts. The remaining 10 watts will be dissipated by the tube.

In a Class B amplifier we will get much better efficiency than in a Class A amplifier. The efficiency will be between 50 and 60 per cent. Let's see what power we can get out of the same tube as in the preceding example in a Class B stage with 50 per cent efficiency.

In the Class B stage, the total power input to the stage must not exceed:

$$10 \times \frac{100}{50} = 20 \text{ watts}$$

This means that with an efficiency of 50 per cent, the power input will be 20 watts and the useful power output 10 watts. The remaining 10 watts will be dissipated by the tube. Notice that we have over twice the power output that we got from a Class A amplifier, and that the plate dissipation rating of 10 watts has not been exceeded.

Let's go one step farther and see what would happen if we used the same tube in a Class C amplifier with an efficiency of 75 per cent. Here the maximum power input to the stage can be:

$$10 \times \frac{100}{25} = 40 \text{ watts}$$

This means that with an efficiency of 75 percent, the power input to the stage can be 40 watts and the useful power output 30 watts. The remaining 10 watts will be dissipated by the tube.

Notice how the efficiency of the amplifier improved as we went from a Class A amplifier to a Class B amplifier. We got a higher per cent of the input power out as useful output power. You can see a still further improvement in going to a Class C stage. Another point that you should notice is the increase in total power that can be handled by a tube in going from a Class A stage to a Class C stage. In a Class A stage, the power input could be a maximum of about 14 watts, but in a Class C stage we can feed 40 watts to the stage. Thus, the efficiency of the stage and the plate dissipation rating of the tube determine the permissible power input to the stage. The input must be limited so that the power wasted in the stage does not exceed the plate dissipation rating of the tube. The power wasted will be determined by the input power and the efficiency of the stage.

SUMMARY

In this section of this lesson we have covered a number of different types of circuits. We do not expect you to remember all the details of each type of circuit at this time. The important thing for you to remember is that there are two types of amplifiers, voltage amplifiers and power amplifiers. Remember also the general appearance of the different circuits. The best way to remember these circuits is by actually drawing them. Notice the similarity between the different types of circuits. In the input of each stage, there is a means of applying the signal between the control grid and cathode of the tube. In each stage, there is some method of developing the required bias. In addition, in each stage you will find some type of load in the plate circuit. By carefully studying these different circuits you will see that there is a great deal of similarity between them and that each circuit is in fact like the basic circuit shown in Fig. 1, but modified for the particular application for which it is designed.

You will study all these circuits in more detail later on, but if you can learn what the circuit looks like now and in general how the amplifier works, you will find it much easier to pick up the various circuit details later.

SELF-TEST QUESTIONS

(l) Into what two types of amplifiers can we divide audio amplifiers?

(m) Try to draw from memory a schematic diagram of a resistance-coupled voltage amplifier.

(n) How is the grid-bias battery eliminated in a typical voltage amplifier?

(o) What limits the size of the resistor that can be used as the grid leak in a voltage-coupled amplifier?

(p) What is the purpose of the bypass capacitor connected across the cathode-bias resistor?

(q) In the power amplifier circuit shown in Fig. 8A, is the output transformer a step-up transformer or a step-down transformer?

(r) Fill in the missing words: a power amplifier using one tube such as shown in Fig. 8 is called a ____________ ended stage whereas one using two tubes such as shown in Fig. 9 is called a __________ ended stage.

(s) What type of amplifier, a voltage amplifier or a power amplifier, is used as the rf amplifier in a radio or TV receiver?

(t) What is the purpose of the capacitor C3 in the rf power amplifier shown in Fig. 11?

(u) Which class of power amplifier has the best efficiency? Which class has the poorest efficiency?

Detectors and Rectifiers

Tubes were used for many years as detectors and rectifiers. However, they are not used for this purpose in modern equipment nearly as often as solid-state devices are. You will study solid-state detectors and rectifiers shortly, but now you will study tube detectors and rectifiers. You will find some tubes used as rectifiers even in modern equipment and, of course, you will probably service many pieces of equipment where tubes have been used as both detectors and rectifiers.

We will study detectors and rectifiers together, because the detector is basically a rectifier. They both work on the principle of allowing current to flow in only one direction. There are, however, some differences in their application. A rectifier is used to change alternating current to direct current. A detector is used to extract information from a radio-frequency carrier.

RECTIFIERS

There are a number of different types of rectifier circuits found in electronic equipment. We will look into two of these circuits in some detail and see in general how these circuits work. There are many details that you will study later.

Half-Wave Rectifiers.

A half-wave rectifier is one that rectifies only half of the ac power-line cycle. During one half cycle, one terminal of a generator is positive and the other terminal is negative. During the next half cycle the terminal that was originally positive becomes negative, and the terminal that was originally negative becomes positive. In a half-wave rectifier circuit the rectifier is arranged so that it conducts when one of these generator terminals is positive but does not conduct when this terminal becomes negative.

A schematic diagram of a half-wave rectifier along with a filter network is shown in Fig. 12. This type of rectifier and power supply is called a universal ac-dc power supply. It is the type of power supply that was used for many years in almost all of the table-model radio receivers manufactured. Even though it is seldom used in modern receivers, the chances are that receivers using this type of supply will be around for many years, and you'll be called on to service a set using this type of power supply. It's called an ac-dc power supply because when it is used in a receiver the receiver can be used on either ac or dc power.

During ac operation, when terminal 1 is positive with respect to terminal 2, the plate of the tube is positive. Thus electrons will be at-

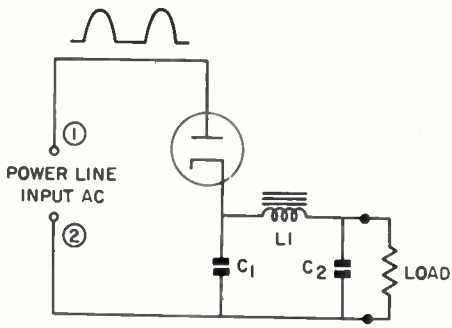


Fig. 12. A half-wave rectifier circuit and a filter network.

tracted from the cathode of the tube to the plate. The complete path for electrons is from terminal 2 through the load, through the filter choke L1, to the cathode of the tube and then through the tube to the plate and back to the other side of the power line. When the polarity of the power line reverses, the plate of the tube becomes negative, and there will be no current flow through the tube and hence no current drawn from the power line.

In a half-wave rectifier, current flows from the power line in a series of pulses as shown in the figure. When the power-line voltage becomes greater than the voltage stored in capacitor C1, there is a large pulse of current that flows through the rectifier tube. This charges capacitor C1. During the remainder of the cycle no current flows through the rectifier tube, and capacitor C1 can be considered as supplying power to the circuit during this part of the cycle. The action of filter choke L1 and filter capacitor C2 is to help smooth the pulsating current to pure dc. The action of these components will be treated in a later text.

Although tube type half-wave rectifiers of the tube just studied are not used in modern receivers, very similar types of rectifier circuits using tubes are still widely used as high-voltage rectifiers in both black and white and color TV receivers and in a circuit called the damper stage in TV. When you study these circuits later in your course, you will see that they are both forms of the half-wave rectifier circuit.

Full-Wave Rectifiers.

The schematic diagram of a full-wave rectifier is shown in Fig. 13. This type of circuit was widely used in both radio and television receivers for many years. Therefore, even though it has been replaced by power supplies using solid-state rectifiers, you should know how it works because you can be sure you will run into power supplies of this type.

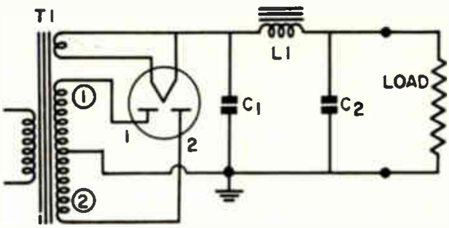


Fig. 13. A full-wave rectifier circuit and a filter network.

In this type of circuit, the high-voltage winding on the power transformer is center tapped, and the center tap is connected to ground. One end of the high-voltage winding is connected to one plate of the rectifier tube, and the other end is connected to the other plate of the rectifier tube. Insofar as the tube operation is concerned, it acts like two separate diodes. When the terminal marked 1 on the high-voltage secondary is positive with respect to the center tap, terminal 2 will be negative with respect to the center tap. Thus the rectifier plate that is connected to terminal 1 will be positive and the rectifier plate connected to terminal 2 will be negative. During this half cycle, current flows from the center tap of the high-voltage winding on the transformer, through the load, through the filter choke to the filament of the rectifier tube. Current then flows from the filament to plate 1 to terminal 1 of the power transformer.

During the next half cycle, terminal 1 of the transformer will be negative with respect to the center tap, and terminal 2 will be positive. During this half cycle, current flows from the center tap through the load, through the filter choke, to the filament of the rectifier tube, to plate 2 of the rectifier tube and then to terminal 2 of the transformer.

As you can see, in the full-wave rectifier circuit of this type current flows first through one half of the high-voltage secondary winding and one half of the tube and then through the other half of the high-voltage secondary winding and the other half of the tube. Thus, if the power supply is operated from a 60-cycle power line there will be two current pulses each cycle. This means that there will be 120 pulses available to charge the filter capacitor marked C1. This is twice as many pulses as can be obtained from the half-wave rectifier. The greater the number of pulses, the easier it is to obtain pure dc. Therefore, it is much easier to filter the pulsating dc at the output of a full-wave rectifier than at the output of a half-wave rectifier.

Pulsating dc at the output of the rectifier is often referred to as dc with a ripple voltage or hum voltage superimposed on it. You will see later that there will always be some ripple present at the output of the power supply regardless of how effective a filter network is. The filter network is designed to cut the ripple to such a low value that it does not appreciably affect the performance of the equipment. The filter systems in most power supplies reduce the ripple voltage so that it is 1% or less of the dc output voltage.

DETECTORS

There have been a number of different kinds of detectors used in

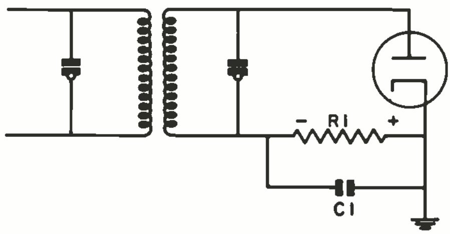


Fig. 14. A diode detector circuit.

electronic equipment, but most pieces of modern electronic equipment use a diode detector. A schematic diagram of a diode detector using a vacuum tube is shown in Fig. 14.

The diode detector shown in Fig. 14 works in much the same way as the half-wave rectifier. When the plate of the tube is positive, current flows through the tube; when it is negative, current cannot flow.

To see how the diode detector can extract intelligence from a radio-frequency signal, let's look at the modulated rf carrier shown in Fig. 15A. We see here an example of amplitude-modulated radio frequency signals. The amplitude or strength of the radio-frequency signal is varying at an audio rate. The audio signal is the intelligence being transmitted.

In the diode detector circuit the current that flows will depend upon the strength of the rf signal applied to it. Thus, as the strength of the rf signal varies, the strength of the pulses of current flowing through the tube will vary. The current flowing through the tube will flow from the cathode to the plate of the tube, through the secondary of the i-f transformer and then through the load resistor R1, setting up a voltage drop across this resistor having the polarity shown on the diagram. The voltage across this resistor will charge capacitor C1. As the strength of the signal varies, causing the current to vary, the voltage across R1 and hence the charge across C1 will vary. Since the strength of the rf signal is varying at an audio rate, the voltage across R1 will vary at an audio rate. This will cause a voltage to appear across R1 and C1 like the voltage shown in Fig. 15B. Notice that this is the audio signal that was actually being carried by the rf signal.

It might be well to point out at this time that the current flows through the diode in a series of pulses like those shown in Fig. 15C. Notice that these pulses are somewhat similar to the pulses obtained from a half-wave rectifier with the exception that the amplitude of these pulses is varying. These pulses tend to charge capacitor C1. The charge across the capacitor will be determined by the

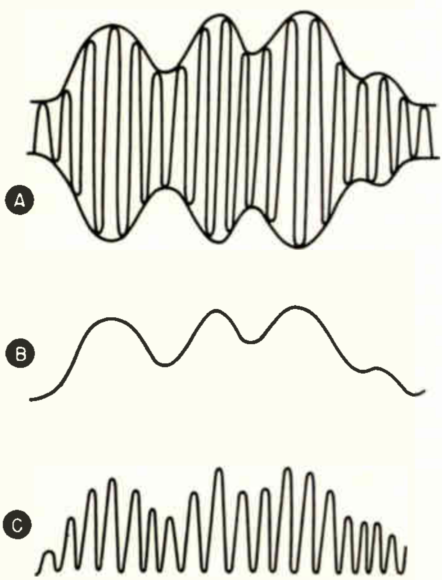


Fig. 15. A modulated rf carrier is shown at A, the extracted audio signal is shown at B, and the series of current pulses flowing through the diode are shown at C.

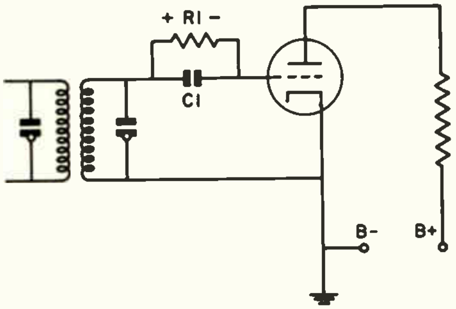


Fig. 16. A grid-leak detector.

amplitude of the pulses, which in turn is determined by the strength of the signal being received and the amplitude modulation on the signal.

In the interval between each cycle there is some tendency for the capacitor to discharge through the resistor, but if the values of the capacitor and the resistor are selected correctly, the discharging is so small that it does not cause any appreciable difficulty. On the other hand, C1 and R1 must be selected so that they are not too large, because C1 must be able to discharge rapidly enough to follow the audio signal variations.

Triode Detectors.

A diode is not the only tube that can be used as a detector. One of the earliest forms of detector was the triode grid-leak detector. The schematic of a grid-leak detector is shown in Fig. 16. The operation of this type of detector is not a great deal different from the operation of a diode detector followed by a triode audio amplifier. If for the present we ignore the plate of the tube and consider only the grid and the cathode, you will see that the circuit is very similar to that of a diode detector. All we would need to do is move the resistor and capacitor to the other side of the transformer to give us the identical circuit.

In a grid-leak detector the rectification actually occurs in the grid circuit with the grid acting like a diode plate. When the rf signal drives the grid positive, electrons are attracted to the grid and current flows to charge the grid capacitor C1. When the rf signal swings negative and there is no grid current flow, capacitor C1 discharges through grid resistor R1, setting up a voltage drop across this resistor as shown. The net result is that we have an audio signal voltage appearing across the grid capacitor and grid resistor. This will cause the grid potential to vary at an audio rate so the tube now acts like a triode audio amplifier and amplifies this audio signal.

The grid-leak detector was quite widely used in the early days of radio, but it has disappeared almost entirely in favor of the diode detector. However, as you can see, the operation is similar to that of a diode detector followed by a triode audio amplifier.

SUMMARY

In this section of the lesson you have seen examples of both detectors and rectifiers. You have learned that a detector is a rectifier inasmuch as it operates on the principle of allowing current to flow through it in only one direction. There is a great deal of similarity between the operation of a detector and that of a rectifier. However, a rectifier is used to convert ac power to dc power where a detector is designed primarily to extract information from an rf carrier signal.

The half-wave rectifier operates on only one half cycle, whereas the full-wave rectifier operates on both half cycles, producing two pulses for each ac power-line cycle.

The diode detector is the most widely used detector in modern electronic equipment. As the amplitude of the signal applied to the diode detector varies, the current flowing through it varies. This causes the voltage across the diode load resistor and capacitor to vary at a rate that follows the intelligence superimposed on the rf signal.

The grid-leak detector is an example of a triode type detector. Its operation is similar to that of a diode detector followed by a triode audio amplifier.

SELF-TEST QUESTIONS

(v) On what principle do detectors and rectifiers operate?

(w) If a half-wave rectifier circuit is operated from a 60-cycle power line, how many current pulses per second will be fed to the filter network?

(x) How many pulses per second will the filter network receive from a full-wave rectifier circuit operating on a 60-cycle power line?

(y) What is the advantage of a full-wave rectifier over a half-wave rectifier?

(z) Draw a schematic diagram of a diode detector circuit.

(aa) What causes the amplitude of the pulses flowing through a diode detector to vary?

(ab) In the grid-leak detector, where does rectification occur?

(ac) To what type of circuit can we compare the grid-leak detector?

Oscillators

One of the most important uses of vacuum tubes is in oscillator circuits. An oscillator is a stage that generates its own signal. Without oscillators, the entire field of electronics would be extremely limited.

In communications, vacuum tubes are used in oscillator circuits to generate radio-frequency signals. Although it is possible to generate ac by mechanical means, there is a limit to how high a frequency can be generated. To generate signals of a very high frequency, electronic means rather than mechanical means must be used.

Practically all oscillators work on the same basic principle: part of the signal from the output of the stage is fed back to the input. The signal fed back to the input is called a feedback signal, or simply feedback. The feedback must be in phase with the signal in the input in order to reinforce it, so the stage can generate an ac signal. Actually, this stage simply converts the dc supplied by the power supply to ac. The exact frequency of the ac signal depends upon the circuit and the value of the components used in the circuit.

HARTLEY OSCILLATORS

Fig. 17 shows an oscillator circuit called a Hartley oscillator. In this circuit, energy is fed from the plate of the tube through C2 to coil L1. This energy is fed through the lower half of the coil to ground. The energy, in flowing through the lower half of L1, sets up a magnetic field which cuts the turns of the upper half of L1 and induces a voltage in it. This voltage is applied between the grid of the tube and ground and produces additional current flow in the plate circuit which will set up a field that will reinforce the signal in the grid circuit still further.

The coil that is marked RFC and is located in the plate circuit of the tube is a radio-frequency choke. It is put in the plate circuit to act as a high impedance to the flow of signal current and thus keep signal currents out of the power supply and force them through capacitor C2 to coil L1.

One of the interesting characteristics of an oscillator circuit is that it develops its own bias. The energy fed from the plate into L1 is of sufficient magnitude to produce a strong enough field to induce a high enough voltage in the upper half of the coil to drive the grid of the tube positive. When this happens, the grid will attract electrons; these electrons will charge the grid capacitor C1 with the polarity shown on the diagram. During the rest of the cycle, when the grid is negative, capacitor C1 discharges through grid resistor R1 and sets up a voltage drop across this resistor, as shown in Fig. 17.

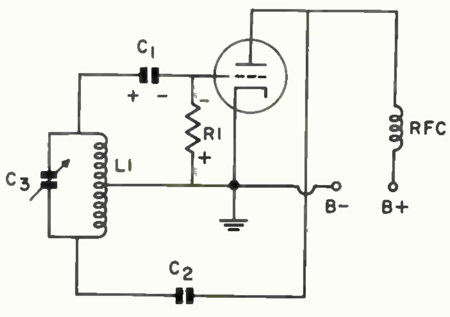


Fig. 17. A Hartley oscillator circuit.

Thus, the grid of the tube is maintained at a negative potential with respect to the cathode. If the amount of feedback from the plate circuit to the grid circuit is increased, the grid will be driven even more positive, charging capacitor C1 still higher, which will increase the bias on the grid of the tube. This in turn will automatically tend to reduce the plate current flowing through the tube. This automatic action tends to adjust the plate current of the tube and maintain it at a nearly constant value.

The frequency at which the circuit will oscillate will be determined primarily by the inductance of L1 and the capacity of C3. This coil and capacitor form a parallel resonant circuit. Changing the value of these components will change the frequency of oscillation.

COLPITTS OSCILLATORS

Another type of oscillator, the Colpitts oscillator, is shown in Fig. 18. Here the feedback is controlled by a capacitive voltage-divider network consisting of C3 and C4. Bias for this stage is produced by grid capacitor C1 and grid resistor R1 as in the Hartley oscillator.

The frequency at which this circuit will oscillate depends primarily on the inductance of coil L1 and the capacity of capacitors C3, C4, and C5. Capacitor C5 is variable; the frequency of the oscillator can be adjusted by changing the capacity of C5. Increasing the capacity will cause the oscillator to operate at a lower frequency; decreasing the capacity will cause the oscillator to operate at a higher frequency.

The output of the Colpitts oscillator as well as that of the Hartley

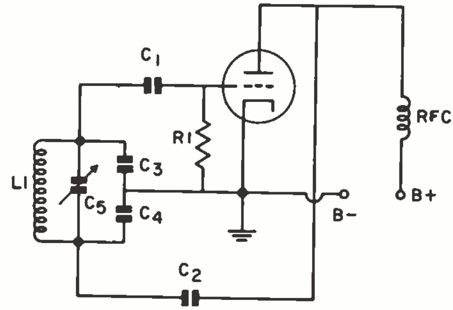


Fig. 18. A Colpitts oscillator circuit.

oscillator is a sine wave. However, not all oscillators have sine-wave outputs. In some cases an output signal other than a sine wave is desired. An example of an oscillator that produces a signal other than a sine wave is the multivibrator.

MULTIVIBRATORS

The multivibrator is widely used in television and in many industrial applications. The output from this type of oscillator is not a sine wave, in fact it is almost a square wave.

The schematic diagram of a typical multivibrator is shown in Fig. 19. This type of multivibrator is called a plate-coupled multivibrator

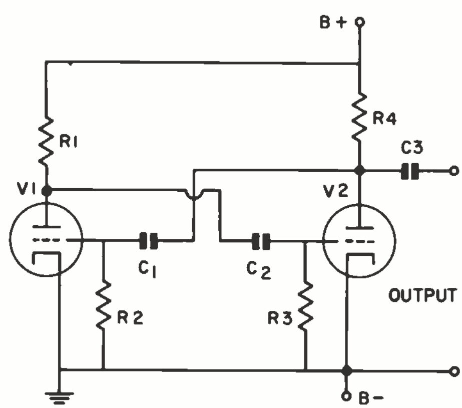


Fig. 19. A plate-coupled multivibrator.

because the energy necessary for oscillation is fed between the tubes from the plate of each tube to the grid of the other tube.

The multivibrator is a rather interesting circuit. Let us see how it works. When the multivibrator shown in Fig. 19 is first turned on, the cathodes of the tubes will not immediately be hot enough to emit electrons, so no current will flow through either tube. Meanwhile the power supply will heat and begin to operate so there will be B+ voltage available. Since there is no current flow through the tubes, there will be no voltage drop across R1 or R4 due to plate current flowing through them. Therefore C1 and C2 will start to charge to a voltage equal to the full power-supply voltage.

To understand this action better, look at Fig. 20. We have shown just the parts involved. The parts shown at A could be drawn as in B; and those shown at C could be drawn as at D. As you can see, when the B supply is operating, C1 will charge through R2 and R4; and C2 will charge through R3 and R1. As these capacitors are charging, the tubes V1 and V2 will be heating and will start passing current.

When V1 and V2 start to pass current, there will be plate current flow through R1 and R4 that will result in voltage drops across these resistors. If C1 and C2 have charged to a voltage higher than the voltage between the plates of the tubes and ground, they must now discharge through R2 and R3 respectively. This will result in a voltage across these resistors that will make the grid end of the resistors negative. The exact voltage will depend on how big a charge the capacitors must get rid of. This in turn will depend on how much current flows through the tubes. It is very unlikely that the two tubes will pass exactly the same current, so one of the capacitors will be discharging at a faster rate than the other, and this will result in a higher negative voltage appearing across one grid resistor than the other.

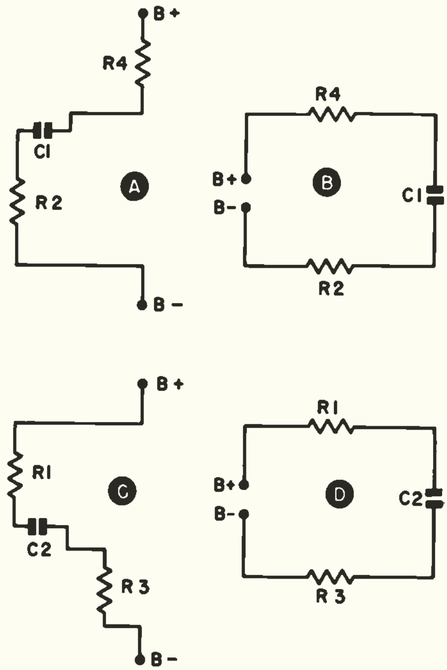


Fig. 20. When the multivibrator shown in Fig. 19 is first turned on, C1 and C2 will charge. This part of the circuit is shown here.

Let's assume that the current through V1 is higher than the current through V2. This will mean that there will be a higher voltage drop across R1 than across R4. C2 will be discharging at a fairly high rate through R3, which will produce a high negative voltage on the grid of V2. This negative voltage on the grid of V2 will reduce the plate current through this tube still further, so the voltage drop across R4 will decrease. C1 will now start to charge to the higher voltage between the plate of V2 and ground. To charge, the capacitor will draw electrons through R2. These electrons will flow through R2 in a direction that will make the grid end of this resistor positive. This positive voltage on the grid of V1 will cause V1 to draw still more current. The higher current through V1 will result in more current flowing through R1, which will cause a greater voltage drop across the resistor. The voltage between the plate of V1 and ground will therefore drop, and capacitor C2 will have to discharge still further. In discharging it will make the grid of V2 even more negative, so the flow of plate current through this tube will be completely cut off.

With plate current through V2 cut off, C1 will eventually charge up to a voltage equal to the B supply voltage. When the capacitor is charged to this voltage, the charging current through R2 will stop flowing, and the positive voltage on the grid of V1 will disappear. When this happens, the current flowing through V1 will drop, causing the voltage drop across R1 to decrease. This will mean that the voltage between the plate of V1 and ground will increase. As soon as this happens, C2 will begin to charge through R3, putting a positive voltage on the grid of V2. Almost instantly V2 will start conducting heavily; the plate voltage on V2 will drop, and capacitor C1 will start to discharge through R2, placing a high negative voltage on the grid of V1 and cutting off the flow of plate current through this tube. When this happens, current will stop flowing through R1, so

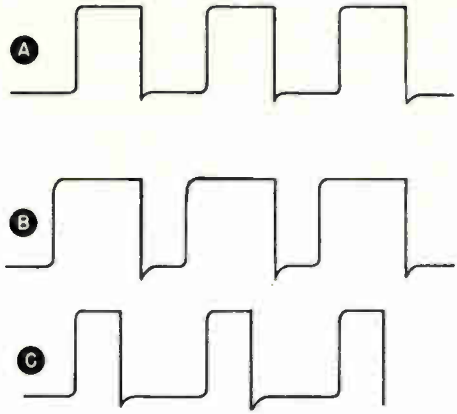


Fig. 21. Output of symmetrical and nonsymmetrical multivibrators.

there will be no voltage drop across this resistor and the voltage between the plate of V1 and ground will jump up to a value equal to the supply voltage. C2 will now have to charge to an even higher voltage, so the grid of V2 will be driven highly positive by the current flowing through R3 to charge this capacitor.

This action of first one tube conducting and cutting off the flow of plate current through the other, and then reversing so that the other tube conducts and cuts off the first tube, will continue as long as power is applied to the oscillator.

Notice that in the multivibrator two tubes are used, whereas in the oscillators we studied before only one tube was needed. In actual practice a dual triode tube that has two separate triodes in the one glass envelope is often used in a multivibrator circuit. Even though this may look like only one tube, we have two-tube action.

The frequency of oscillation will depend primarily on the values of C1-R2 and C2-R3.

If the combination of R2 and C1 has a time constant equal to that of

R3 and C2, the multivibrator is called a symmetrical multivibrator. Each tube will conduct and be cut off for the same length of time, and the output will look like Fig. 21A. However, if the time constant of C2 and R3 is longer than that of C1 and R2, V2 will be cut off for a longer time than V1, and the output will be like Fig. 21B. On the other hand, if the time constant of C1 and R2 is longer than that of C2 and R3, V1 will be cut off longer than V2, and the output will be like Fig. 21C.

The plate-coupled multivibrator is only one type of multivibrator; there are several other types that you will study later.

SUMMARY

Oscillators are important to the electronics technician. You will find them in radio and TV receivers. Every superheterodyne has a local oscillator. Television receivers have oscillators similar to the multivibrator to generate the signals that move the electron beam over the face of the picture tube.

The Hartley oscillator and the Colpitts oscillator both generate a sine-wave output. In the Hartley oscillator, feedback is obtained by inductive means, whereas in a Colpitts oscillator feedback is obtained by means of a capacitive voltage divider. The two oscillators are otherwise basically similar.

The multivibrator is an RC coupled oscillator. Its output is not a sine wave. It is useful in TV receivers and in many industrial applications. Two tubes are needed in a multivibrator circuit. The two conduct alternately; when the first tube is conducting, the second is cut off, and when the second tube is conducting, the first is cut off.

SELF-TEST QUESTIONS

(ad) On what principle do oscillators operate?

(ae) Draw a schematic diagram of a Hartley oscillator.

(af) Across what part is grid bias for the Hartley oscillator developed?

(ag) In the Hartley oscillator circuit shown in Fig. 17, which two parts primarily control the oscillator frequency?

(ah) What controls the feedback in the Colpitts oscillator circuit shown in Fig. 18?

(ai) What parts primarily determine the oscillator frequency in the Colpitts oscillator shown in Fig. 18?

(aj) What type of output signal is obtained from the Hartley and Colpitts oscillators?

(ak) What type of output signal is obtained from a multivibrator?

(al) What parts primarily control the frequency of the multivibrator shown in Fig. 19?

(am) What is the name given to the particular multivibrator shown in Fig. 19?

A Complete Superheterodyne Receiver

Now let us see how the different circuits are put together in a radio receiver. We have already mentioned that modern radio receivers use what is called the superheterodyne circuit. A block diagram of a typical superheterodyne receiver is shown in Fig. 22. In Fig. 22A we have shown the various functions performed in the receiver. The block diagram is often drawn like the diagram in B because the mixer and oscillator are usually combined in one tube and the second detector and first audio stage are also combined in the tube. Therefore the diagram, shown in Fig. 22B, shows in block diagram form the functions performed by the various tubes.

In the superheterodyne receiver, the signal is picked up by the antenna and is fed to the first stage which is the mixer stage as shown in Fig. 22A. At the same time, a signal from a local oscillator is also fed to the mixer stage. The local oscillator always operates at a fixed frequency above the frequency to which the mixer is tuned. In modern radio receivers the oscillator is usually operated 455 kHz above the incoming signal. In the mixer circuit the in-

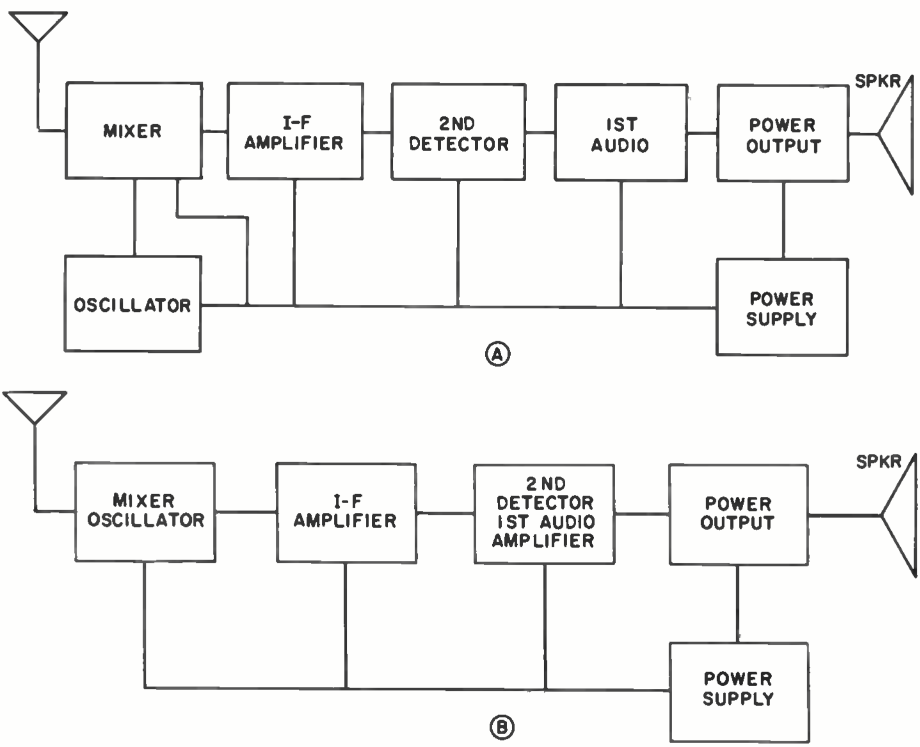


Fig. 22. A block diagram of a 5-tube superheterodyne receiver.

coming signal is mixed with the signal produced by the local oscillator. This mixing of the two signals results in two new signal frequencies appearing in the output of the mixer, one equal to the sum of the local oscillator and the incoming signal frequency, and the other equal to the difference between the two frequencies.

We mentioned that the oscillator is operated 455 kHz above the incoming signal frequency. Therefore the difference between the incoming signal frequency and the oscillator signal frequency will be 455 kHz.

The second stage in the superheterodyne is a radio-frequency amplifier stage called the i-f amplifier. It is tuned to the i-f frequency. which is equal to the difference between the frequency of the incoming signal and the frequency of the local oscillator in the receiver.

The i-f amplifier is what is called a fixed-frequency amplifier. In other words, it is always tuned to the same frequency, which in most cases is 455 kHz. Because the amplifier is tuned to a fixed frequency, it is possible to design a stage with a very high gain without running into any problems such as instability or oscillation. Therefore the i-f amplifier amplifies the i-f signal substantially.

The i-f signal is then fed to the stage called the second detector, which is usually a diode detector. Here the intelligence is separated from the radio-frequency carrier. The intelligence signal is the audio signal used to modulate the carrier signal at the transmitter. The audio signal produced by the second detector is fed to the first audio stage where it is amplified and finally fed to the output stage, which is a power amplifier that produces the power necessary to drive the loudspeaker.

A schematic diagram of a superheterodyne receiver is shown in Fig. 23. We will not go through this receiver stage by stage to see how each stage works. We have already studied most of these stages in this lesson, so we will now concentrate on taking up a few additional details and also seeing how the stages are used together.

THE MIXER-OSCILLATOR

The tube used in the mixer-oscillator stage is a 12BE6 tube marked V1. This tube is a pentagrid converter. Pentagrid means five grids; the tube is called a converter because it is designed for service as a frequency converter, which is what the mixer-oscillator stage is often called.

In this tube the first and second grids in conjunction with the cathode act like a triode tube. The first grid is the control grid and the second grid acts as the plate of the triode section of the tube. Thus the cathode, the first grid, and the second grid are used as an oscillator. This oscillator modulates the stream of electrons flowing from the cathode to the plate of the tube. The incoming signal picked up by the loop antenna is fed into the No. 3 grid. This signal also modulates the electrons flowing from the cathode to the plate of the tube so that the resultant flow of electrons from the cathode to the plate of the tube will be modulated both by the signal produced by the local oscillator and by the incoming signal. This modulation will produce a signal with a frequency equal to the difference in the frequency of the

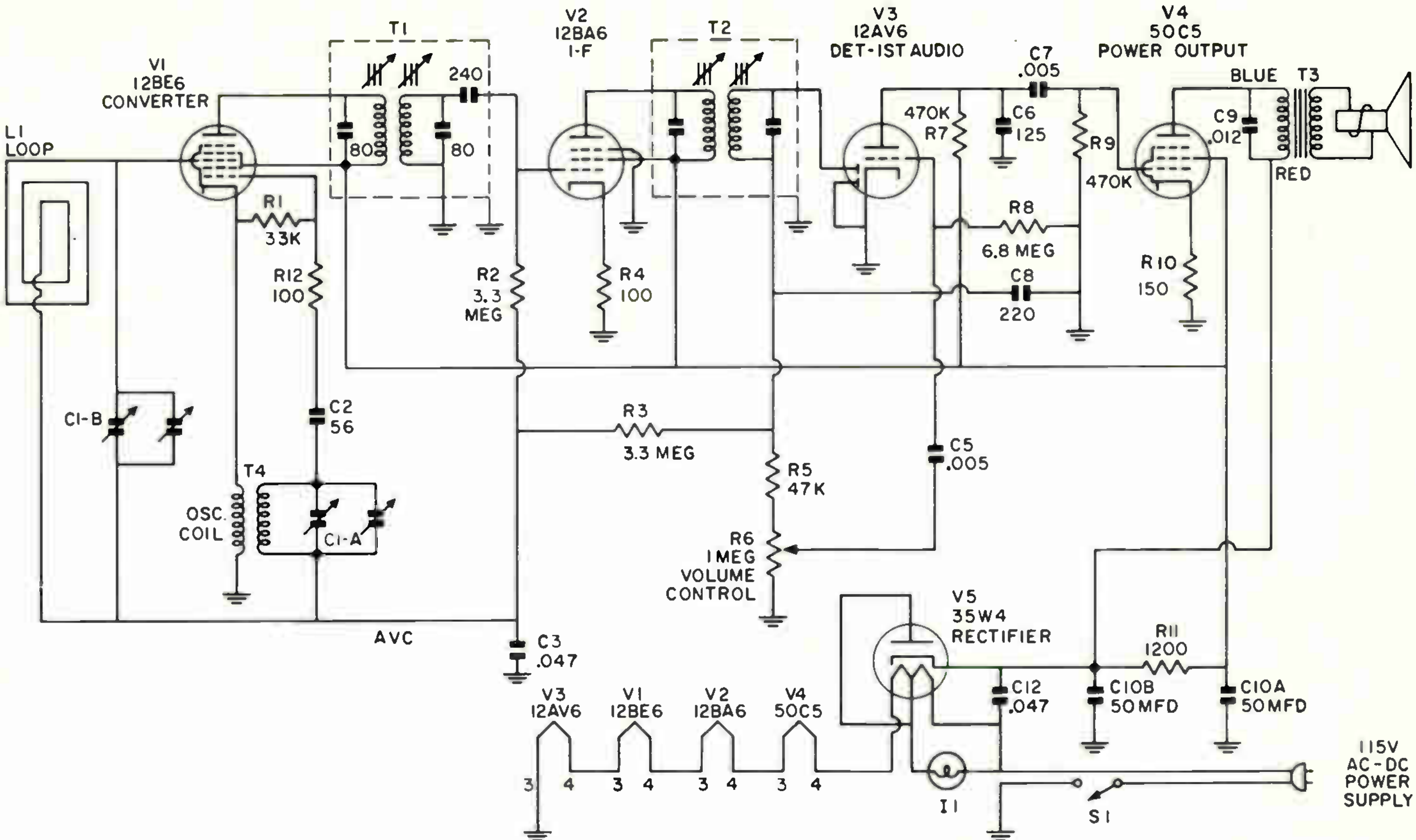


Fig. 23. Schematic diagram of a complete superheterodyne receiver.

two signals which will be the i-f signal, in addition to a signal equal to the sum of the frequencies of the two signals.

The primary winding of the transformer T1 and the capacitor connected across it form a parallel resonant circuit. This circuit is tuned to resonance at the difference frequency produced in the converter stage. Therefore it acts as a high resistance at this frequency, while at the same time it acts as a low impedance at the frequency of the incoming signal, the frequency of the local oscillator, and the frequency equal to the sum of the two frequencies. Thus in the parallel resonant circuit we will have a high circulating current having a frequency equal to the difference frequency. The other three frequencies appearing in the plate circuit of the mixer tube will produce little or no current flow in the resonant circuit.

THE I-F AMPLIFIERS

The secondary of T1 is inductively coupled to the primary. Thus a voltage will be induced in series with the secondary winding. The secondary winding of the transformer is tuned to resonance by the capacitor connected across it, and this coil and capacitor form a series resonant circuit. In a series resonant circuit there will be a high circulating current, and a resonant voltage step-up. This voltage is applied between the grid and cathode of the i-f amplifier.

The difference-frequency signal applied between the grid and cathode of the i-f amplifier is amplified by this stage. In the plate circuit of this tube, which is a conventional pentode rf amplifier, there is another parallel resonant circuit. This is made up of the primary winding of T2 plus the capacitor connected across it. Again, this transformer acts as a high resistance at the difference frequency and a low impedance at other frequencies.

Before leaving the i-f stage, notice the lead marked AVC. AVC is an abbreviation for automatic volume control. A variable voltage is applied to this lead, which in turn is connected to the grid of the i-f tube V2 through the 3.3-megohm resistor R2. The voltage is also applied to the grid of the mixer tube through the loop antenna.

The voltage that is fed to the grids of these tubes will depend upon the strength of the signal picked up from the station. If the signal is very strong, a fairly high negative voltage is fed to the grids of these tubes and this voltage reduces the gain of the stages. If the signal is weak, the negative voltage drops almost to zero so that the two stages operate at maximum gain. We will see how this voltage is developed when we study the second detector used in this receiver.

The tube used in the i-f stage of a modern superheterodyne receiver is usually what is called a remote cut-off type. A remote cut-off tube is a tube with a specially made control grid. The grid is designed so that it takes a very high negative voltage to completely cut-off the flow of plate current.

In the tubes we have discussed so far, the grid wires have been evenly and rather closely spaced. The plate current flow in a pentode tube that is made in this way can be cut off with a fairly low negative voltage applied to the grid of the tube. This type of tube is called a sharp cut-off tube. If however, instead of spacing the grid wires evenly, we space them so that they are close together at the ends but spaced quite widely apart in the middle as shown in Fig. 24, a much higher negative voltage must be applied to the grid of the tube before the plate current can be reduced to zero. As we begin to apply negative voltage to the grid of the tube to cut off the flow of plate current, the grid wires at the ends are able to cut off the flow of plate

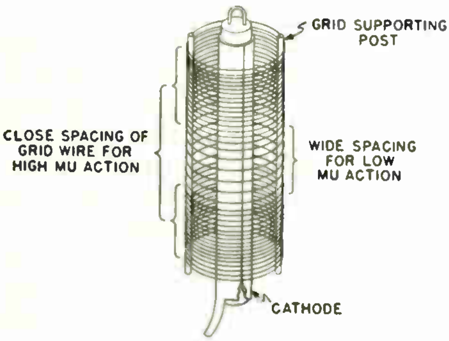


Fig. 24. The grid of a remote cut-off type pentode. Here the spacing between the grid wires is greater at the center of the grid than at the ends.

current through them, but because of the wide spacing in the center part of the grid, electrons will still travel from the cathode to the plate of the tube. In order to cut off the flow of plate current with this type of grid structure, a much higher negative grid voltage is required than with the sharp cut-off tube.

The remote cut-off tube is ideally suited for i-f amplifiers where avc is used. However, this type of tube usually does not have quite as high a gain as a sharp cut-off tube. A compromise between the sharp cut-off and the remote cut-off tubes is the semi-remote cut-off tube. In this type of tube, the spacing between the grid wires in the center of the tube is not quite as wide as the spacing between the grid wires in the remote cut-off tube.

THE DETECTOR - FIRST AUDIO STAGE

In our superheterodyne receiver the i-f signal is fed from the i-f stage to the second detector. The secondary of the i-f transformer forms a series resonant circuit with a capacitor connected across it. One end of the secondary is connected directly to the plate of the diode detector and the other end is connected to the diode load, which is made up of the 47K-ohm resistor R5 and 1-megohm volume control R6. These two resistors are in parallel with the diode load capacitor C8. You already know how a detector works and how an audio voltage will appear across the diode load resistors R5 and R6 and the diode load capacitor C8.

The i-f signal flowing through the detector can flow in only one direction with the result that there will be a series of pulses at the detector output which will charge the diode capacitor C8. The charge across this capacitor will depend upon two things, the audio signal and also the strength of the signal being received. By connecting a filter network consisting of a resistor and a capacitor such as R3 and C3 across this capacitor a dc voltage can be obtained across C3 that will depend upon the strength of the incoming signal. Now let's see where this voltage comes from.

We have a series of pulses flowing through the diode, charging capacitor C8. The amplitude of these pulses, and hence the voltage across C8, depends on the audio signal and the strength of the signal being received. C8 discharges through R5 and R6, producing a voltage having a polarity such that the junction of R5 and the i-f transformer is negative. This voltage also will depend on the audio signal and the strength of the signal being received.

This voltage also is across the combination of R3 and C3 in series. However, R3 and capacitor C3 form a voltage divider network. The resistance of R3 is much higher than the reactance of C3. Therefore most of the audio signal appearing across the two will be dropped across R3 so that the voltage across C3 will be almost pure dc and its strength will depend upon the strength of the incoming signal. This voltage is then fed to the avc line and used to control the gain of the mixer and i-f tubes. If the strength of the signal being picked up is strong, then a fairly high negative voltage is developed across C3, whereas, if the signal is weak, the voltage across C3 will be low. This voltage is used to control the gain of the mixer and i-f stages so they operate at maximum gain when a weak signal is being received, and at reduced gain when a strong signal is being received.

AVC is used in most modern radio receivers to regulate the gain of the set, so that as you tune across the broadcast band from one station to another, they all come in at approximately the same volume. Actually, this system is an automatic gain control rather than an automatic volume control, but when the scheme was first introduced manufacturers called it an automatic volume control because they felt that this would have more appeal to the public than the name automatic gain control. The name automatic volume has stuck, but the same system used in television receivers is called automatic gain control.

The tube used in the second detector-first audio stage is a combination tube. The tube contains two diodes and one triode. Only one of the diodes is used in the detector circuit; the other diode plate is unused and simply connected to ground and to the cathode. The triode section is used as the first audio stage. You will notice that the diode load is actually made up of two resistors. The 47K resistor R5 is only one twentieth of the size of R6 so most of the voltage will appear across R6. R6 is a potentiometer. The position at which the center tap connects to the resistor can be controlled by rotating the control shaft on the potentiometer. By moving this tap up and down the resistor, the amount of audio signal fed to the first audio stage can be varied. This control is called a volume control since it will vary the volume or output from the sound system in the receiver. The audio signal is then fed through coupling capacitor C5 to the grid of the first audio tube.

The first audio tube uses a bias system that we have not discussed previously. Notice that the cathode of the tube is connected directly to ground, and the grid resistor R8 is a 6.8-megohm resistor. You already know that some of the electrons leaving the cathode of the tube and travelling toward the plate will accidentally strike the grid of the tube. The number striking the grid is quite small, and these electrons flow through the grid resistor back to ground and from there back to the cathode of the tube. In most stages the grid resistor is kept low enough so the number of electrons flowing through it do not produce an appreciable voltage. However, by using a large value of grid resistor, a voltage can be developed across it to bias the tube. This type of bias is usually called "convection" bias and is frequently used in the first audio stage of modern radio receivers.

THE OUTPUT STAGE

The output stage in most modern receivers is either a pentode or a beam power tube. The tube is operated as a Class A amplifier.

In this receiver, the signal from the plate of the first audio stage is coupled to the grid of the output stage through capacitor C7. R9 is a grid leak resistor and R10 is the cathode bias resistor. The cathode bypass capacitor is omitted from this stage to improve the frequency response of the stage. You will see in a later lesson exactly why this happens.

The loudspeaker is a permanent-magnet dynamic speaker. This speaker is coupled to the output tube by means of transformer T3, which is called the output transformer. We have already studied output transformers and know that they are impedance-matching devices designed to match the low impedance speaker to the high impedance of the plate of the output tube in order to get as high a power transfer as possible.

There are a number of different types of tubes used as output tubes, but in general the circuit is similar to the one shown in Fig. 23, although the values of the components used in the circuit may vary slightly with different tubes.

THE POWER SUPPLY

The only section of the receiver left to discuss is the power supply. This is a universal ac-dc type of power supply using a half-wave rectifier similar to the one shown in Fig. 12.

The Plate Supply.

Notice that the plate voltage for the output tube is taken directly from the cathode of the rectifier tube through the primary of the output transformer instead of from the output of the filter network. We can do this because the plate current flowing in a pentode or beam power tube depends primarily on the screen voltage, and the plate voltage has little or no effect on it. Thus the ac ripple applied to the plate of the tube will not cause any appreciable hum current to flow through the tube and the primary of the output transformer. The output tube alone will draw as much current as all the rest of the tubes in the receiver. If we take the plate supply of the output tube directly from the cathode of the rectifier, we can use a resistor instead of a choke in the filter network, as we have done here-R11 in Fig. 23.

If we connected the plate of the output tube to the other side of the filter network, we would have to use a choke, because the high current drawn by the output tube through a resistor would cause such a high dc voltage drop that the output voltage for the rest of the plates would be too low. A choke has a high ac reactance, so it would filter any ripple voltage, but it has a low dc resistance, so the dc voltage drop across it would be lower.

The advantage of using a resistor in place of a choke is simply one of economy. A resistor is much cheaper than a filter choke. As long as the current through the resistor can be kept to a reasonably low value, the dc voltage drop across the resistor will not be too great. The resistor will form a voltage divider network with the output filter capacitor so that the hum voltage from the cathode of the rectifier will be reduced enough to provide suitably pure dc for the operation of the receiver. The plate and screen voltages for the rest of the tubes and the screen voltage for the output tube are obtained from the output of the filter network. These voltages will be much better filtered than the voltage supplied to the plate of the output tube. If these tubes, particularly the first audio tube or the screen of the output tube, were operated directly from the cathode of the rectifier tube there would be an extremely loud hum produced by the speaker.

The Heater Circuit.

The heaters of the various tubes in this receiver are connected in series, and the series circuit is connected directly across the power line. The tubes are designed so that they all require the same operating current. The voltage requirements for the various tubes are different. The 35W4 rectifier tube requires a heater voltage of 35 volts. The 50C5 output tube requires a heater voltage of 50 volts. Each of the other tubes requires a heater voltage of 12.6 volts. If you add these voltages together, you will find that the total is 122.8 volts. In actual practice, these tubes can be operated from a line voltage of anywhere between about 110 volts and 125 volts, and give satisfactory performance. Notice the pilot light marked I1 on the diagram. The heater of the 35W4 is tapped, and the pilot light is connected in parallel with part of the heater to get the voltage needed to operate the light.

Notice that 12-volt heaters are closer to the side of the power line that connects to B-, so that the voltage between these heaters and B- is maintained as low as possible.

The first audio tube is almost always connected at the end of the string so that one side of its heater is connected to B-. This keeps the potential between the cathode and heater of this tube as low as possible. This particular tube is more susceptible to hum pickup than are the other tubes in the receiver. The arrangement used in this receiver is pretty standard for this type of set. If you trace the heater circuit in an ac-dc receiver starting at the side of the power line that connects to B-, you will find the tubes connected in the following order: the first audio-second detector; the mixer oscillator; the i-f amplifier; the power output tube, and finally the rectifier.

Since the tube heaters in this receiver are connected in series, if any one of the heaters burns out, none of the tubes will light. This is probably the most common defect you will find in ac-dc receivers. Keep this point in mind. If you are asked to service a small ac-dc receiver and you find that none of the tubes light, look for a tube with an open heater. The chances are that replacing this tube will clear up the trouble.

SUMMARY

In this section of this lesson we have shown how the various stages that you have studied previously are put together to form a complete receiver. From this discussion you can see that a complete radio receiver is simply made up of a number of different stages, each performing the task for which it was designed. Complex electronic equipment is made up in the same way; it is made of a number of simple separate stages designed to work together.

There are a number of things that we have not discussed about these various stages, but we will discuss them in more detail in later lessons. At present you should have a general understanding of how each stage works and how they are put together to form a complete receiver.

SELF-TEST QUESTIONS

(an) To what stage is the signal from the antenna fed in a typical five-tube superheterodyne receiver?

(ao) In a modern superheterodyne receiver is the oscillator frequency higher or lower than the frequency of the incoming signal?

(ap) If you tune a superheterodyne receiver to a broadcast station operating on 1500kc, and the i-f amplifier operates on a frequency of 456kc, at what frequency must the local oscillator in the receiver be operating?

(aq) In the 12BE6 tube used in the circuit shown in Fig. 23, which grid acts as the plate of the oscillator tube?

(ar) In the i-f transformer T1 in Fig. 23, both the primary and secondary circuits are tuned to resonance. Are these circuits series-resonant circuits or are they parallel-resonant circuits?

(as) What is meant by a remote cut-off tube?

(at) In the second detector in the circuit shown in Fig. 23, across what part is the audio signal voltage developed?

(au) How is bias for the first audio stage in the circuit shown in Fig. 23 developed?

(av) What class of power amplifier is used in the circuit shown in Fig. 23?

(aw) In the circuit shown in Fig. 23, the heaters of all the tubes are connected in series across the power line. If you were called on to service a receiver of this type and saw that none of the tubes were lighting, what type of trouble would you look for?

Amplifier Variations

We have already mentioned that there are three basic types of amplifiers, the grounded-cathode, the grounded-grid, and the grounded-plate. All the amplifiers you have studied so far have been grounded-cathode amplifiers. Most of the amplifiers you will encounter will be the grounded-cathode amplifiers, but the other two types are also important. They have some special characteristics that are very useful in certain applications. The grounded-grid amplifier is widely used as the rf amplifier in VHF tuners of TV receivers. It is also quite widely used in transmitting equipment. The cathode-follower, which is the more frequently used name for the grounded-plate amplifier, is widely used in industrial applications, and also it has recently been used in a large number of color TV receivers. The cathode-follower is particularly useful as an impedance-matching device to match a high-impedance circuit to a low-impedance circuit.

In this section of this lesson you will study some of the important characteristics of all three types. You will see how these amplifiers differ from each other and see some examples of grounded-grid and grounded-plate amplifier circuits.

THE GROUNDED-CATHODE AMPLIFIER

As we mentioned, the amplifiers we have been studying have all been grounded-cathode amplifiers. We have repeated the basic grounded-cathode circuit shown in Fig. 1 as Fig. 25. Notice that the signal is applied between the grid and cathode of the tube; the cathode is operated at signal ground potential, and the load is placed in the plate circuit. The output signal is developed across the load, so in this circuit it is developed between the plate of the tube and ground. Remember that the end of the load that is connected to the B supply is at ground potential as far as the signal is concerned.

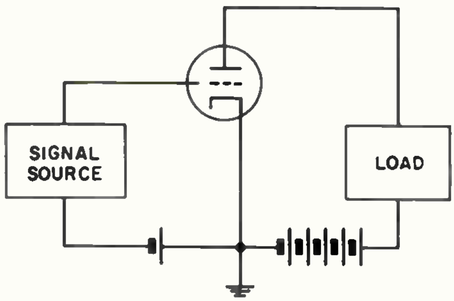


Fig. 25. The grounded-cathode amplifier shown in Fig. 1 is repeated here for your convenience.

Circuit Characteristics.

You already know how the grounded-cathode circuit amplifies a signal, so we will not go through an explanation of how this circuit works. Instead we will discuss some of the important characteristics of this type of circuit.

One important characteristic of this type of amplifier is that the output signal is greater than the input signal; in other words, the stage is capable of amplifying. The actual gain of the stage is equal to the ratio of the output voltage divided by the input voltage. The gain obtained from this type of stage may be quite low, sometimes as low as only two or three, or it may be quite high, sometimes as high as two hundred or more. The exact gain will depend on the tube used in the amplifier and the value of the components used with the tube.

Another important characteristic of this type of stage is that there is a phase shift of a half-cycle or 180° between the input and output signals. This means that when the input signal drives the grid in a positive direction, the output signal between the plate and ground will be going in a negative direction. Similarly if the grid is being driven negative, the output signal will be going positive.

Two other important characteristics are the input and output impedances of the stage. The input impedance is important because it tells you how the stage will load the signal source. The output impedance is important because it gives an indication of the type of load that must be connected in the output circuit in order to obtain proper results from the stage.

You will remember that in a Class A voltage amplifier the tube is biased so that it operates on the mid-point of its characteristic curve. The input signal is not strong enough to drive the grid positive, so there will be no grid current flow. Thus the input circuit of a Class A voltage amplifier is a high-impedance circuit. This is important because it tells us that this type of stage will take little or no power from the source driving the stage.

In a Class B or Class C power amplifier the grid is driven positive, so grid current will flow. Thus these stages have a comparatively low input impedance and will require power from the driving stage.

We can summarize the input characteristics of the grounded-cathode amplifier as follows: Voltage amplifiers and Class A power amplifiers draw no grid current, and hence have a high input impedance. Class B and Class C power amplifiers draw grid current, require power from the driver stage, and have a fairly low input impedance.

The grounded-cathode amplifier also has a fairly high output impedance. The load is placed in the plate circuit of the tube, and the output is developed across this load. If the load impedance is low, the output will be low.

THE GROUNDED-PLATE AMPLIFIER

The grounded-plate amplifier is more commonly called a cathode follower. The load in this type of circuit is connected in the cathode circuit between the cathode of the tube and ground. A block diagram of this type of amplifier is shown in Fig. 26.

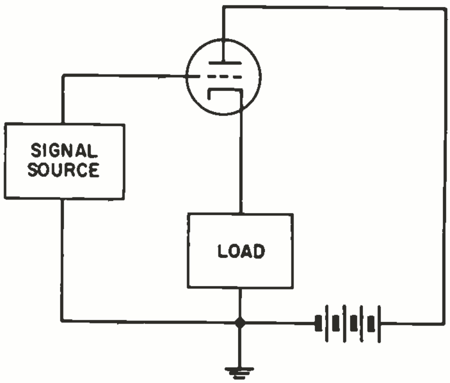


Fig. 26. A grounded-plate or cathode-follower circuit.

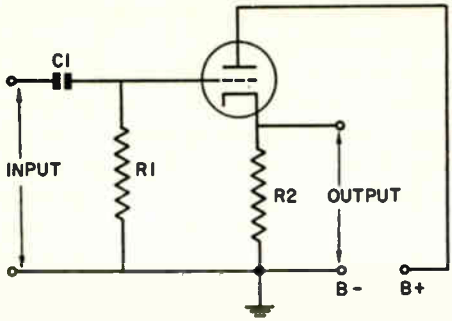


Fig. 27. Schematic diagram of a cathode follower stage.

Notice that the signal is applied between the grid of the tube and ground, as in the grounded-cathode amplifier, but the difference between the two is in the placement of the load.

Putting the load in the cathode circuit completely alters the characteristics of the amplifier. Let us study the operation of a typical cathode follower such as the one shown in schematic form in Fig. 27 to see why this type of amplifier is so different from the grounded-cathode type.

Capacitor C1 is used in the input circuit to isolate the grid circuit of the stage from the signal source and to keep any dc that may be present at the signal source away from the grid of the tube. Resistor R1 is the grid resistor and is placed in the grid circuit to act as a grid leak and provide a path back to the cathode for any electrons that accidentally strike the grid of the tube. The load is R2. It is placed in the cathode circuit of the tube between the cathode and B-. The plate of the tube is connected directly to B+, and since B+ is at ground potential insofar as the signal is concerned, the amplifier is a grounded-plate amplifier.

The operation of this type of amplifier is quite different from that of the grounded-cathode amplifier. When the input signal drives the grid in a positive direction, the current flowing from the cathode of the tube to the plate will increase. This will cause the voltage across the cathode resistor R2 to increase because the plate current flows through this resistor. However, notice that the input signal is applied between the grid of the tube and ground. The actual voltage that will control the flow of plate current through the tube is the voltage between the grid and cathode. If the grid voltage goes positive, the increase in plate current will cause the voltage between the cathode and ground to increase and become more positive. In other words, the positive voltage at the grid produces a more positive voltage at the cathode. This subtracts from the signal voltage so that it reduces the net grid-to-cathode voltage. This means that we have a situation where the output signal being produced by the stage subtracts from the input signal and reduces the input to the tube. Thus the output signal must always be less than the input signal in this type of amplifier. Technicians say that the gain of the stage is less than one. If the output signal were greater than the input signal, it would completely cancel the input signal producing it.

You might wonder why the cathode follower would be of any use at all if the output signal is less than the input. The stage is useful for two reasons. First, the output signal is in phase with the input. This means that when the input signal swings positive, the output signal swings positive; and when the input signal swings negative, the output signal swings negative. There are some applications where it is important not to change the phase.

Another and more important use of the cathode follower is as an impedance-matching device. The cathode follower has a high input impedance, but because the load is in the cathode circuit, it has a low output impedance. Thus, if we have a high-impedance generator and want to connect it to a low-impedance load, we can use a cathode follower to match the two impedances.

Cathode followers are very useful in any application where it is important to isolate the load from the signal source. They are often found in the input of a cathode ray oscilloscope, which is a test instrument used in servicing many kinds of electronic equipment. The cathode follower is used in the input circuit to provide isolation between the circuit under test and the test instrument so that connecting the test instrument to the equipment does not affect the performance of the equipment.

THE GROUNDED-GRID AMPLIFIER

The grounded-grid amplifier is used in some special cases as a power amplifier. You will remember that we pointed out that when a triode tube is used as an rf amplifier, unless special precautions are taken, the tube will oscillate. This oscillation that occurred in a triode tube was one of the factors that led to the development of the tetrode tube. Remember that the screen grid was added to the tube in order to shield the grid from the plate and keep the energy fed from the plate back to the grid of the tube as low as possible.

In some circuits a triode tube will operate where a beam power tube or a pentode tube will not give satisfactory service. This situation is encountered at very high frequencies. It comes about because of the time it takes an electron to travel from the cathode of the tube to the plate. At very high frequencies a single cycle may take only a fraction of a millionth of a second. A half cycle will take only half this time. It could take the electrons longer than the time of one half cycle to travel from the cathode to the plate of the tube if the spacing between the cathode and plate were too great. The spacing between the cathode and plate in a beam power tube or a pentode is greater than in a triode because there must be room for the extra elements.

In higher-power amplifiers using beam power tubes with a high sensitivity, there will sometimes be oscillation because of energy getting through the screen from the plate of the tube to the grid. If the tube is a very sensitive tube, a small amount of energy getting through the screen may be enough to cause oscillation in the stage.

Oscillation in a triode stage or in a high-gain beam power stage can often be eliminated by using a grounded-grid circuit, because the grid is operated at ground potential. The signal is fed into the cathode circuit, and the grid, which is grounded, acts to shield the output circuit from the input circuit.

A simplified diagram of a grounded-grid amplifier is shown in Fig. 28. Notice that in this circuit the input is fed into the stage between the cathode and ground, and the grid

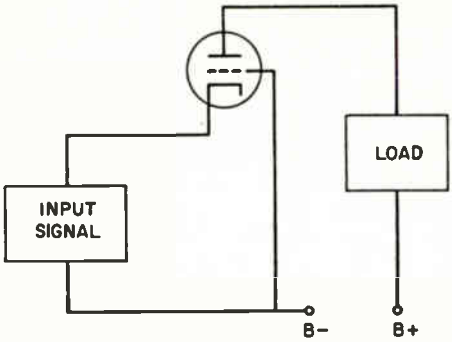


Fig. 28. Block diagram of a grounded-grid amplifier.

is operated at signal ground potential. The load is connected in the plate circuit of the tube as in the grounded-cathode amplifier.

The schematic diagram of a grounded-grid amplifier using a beam power tube is shown in Fig. 29. In this circuit the input signal is fed between the filament-type cathode and ground. The filament of the tube is isolated from ground by the two radio-frequency chokes, marked RFC1 and RFC2 on the diagram. These chokes are simply coils that have high inductive reactances at the operating frequency. Because they have high reactances, the filament cathode is isolated from ground insofar as the signal is concerned, while at the same time the chokes offer little or no opposition to the flow of dc or low-frequency ac used to heat the filament of the tube.

In this type of amplifier, if the incoming signal drives the cathode positive with respect to ground, it drives the cathode positive with respect to the grid, because the grid is connected to ground as far as the signal is concerned through capacitor C1, which has a low reactance at the signal frequency. If the cathode is driven positive with respect to the grid, it is the same as driving the grid negative with respect to the cathode, so the current flowing from the cathode to the plate of the tube will decrease. On the other hand, if the cathode is driven negative with respect to the grid, it is the same as driving the grid positive with respect to the cathode, so the plate current will increase.

Since both the grid and the screen grid of this tube are connected to ground as far as the signal is concerned, it will be almost impossible for any energy to get from the plate of the tube back into the input circuit. As long as energy cannot get from the plate circuit back into the input circuit, the stage will not oscillate.

It is not important to go into all the details of the grounded-grid amplifier at this time. The important thing for you to remember is the general configuration of the circuit. Remember that the signal is fed into the cathode circuit, and the load is connected into the plate circuit. Remember also that the big advantage of this type of circuit is that the energy fed from the plate circuit

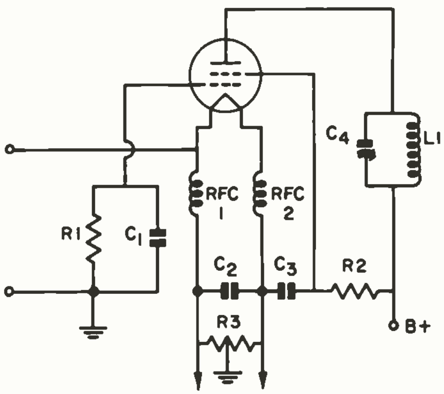


Fig. 29. Schematic diagram of a grounded-grid amplifier.

back into the input is low, so that the possibility of oscillation is remote.

SUMMARY

The important thing for you to remember from this section of the lesson is the general pattern of the circuit. Remember that there are three types of amplifier circuits: the grounded-cathode, the grounded-plate or cathode follower, and the grounded-grid amplifier. Most of the amplifiers that you will encounter will be grounded-cathode amplifiers, but the other types are found in some special applications.

The grounded-cathode amplifier is important because a high gain can be obtained from this type of circuit. This type of amplifier has a high input impedance and a high output impedance.

The cathode follower circuit always has a gain of less than one. This means that the output will be less than the input. This circuit is often used as an impedance-matching device; it has a high input impedance and a low output impedance.

The grounded-grid amplifier is often used as an rf power amplifier. It is particularly useful with triode tubes operating at very high frequencies and with beam power tubes having a high sensitivity. In conventional circuits, the screen may not offer enough isolation to prevent a beam power tube with a high sensitivity from oscillating.

SELF-TEST QUESTIONS

(ax) Where are the signal source and the load placed in a grounded-cathode amplifier?

(ay) Name two important characteristics of the grounded-cathode amplifier?

(az) What is the more common name by which the grounded-plate amplifier is known?

(ba) Where are the signal source and load placed in the cathode follower?

(bb) How does the amplitude of the output signal compare with the amplitude of the input signal in a cathode follower?

(bc) How does the phase of the output signal compare with the phase of the input signal in a cathode follower?

(bd) Name two important applications where a cathode-follower stage may be used?

(be) Where are the input signal and load placed in the grounded-grid amplifier?

(bf) How is oscillation prevented when a triode tube is used as an rf amplifier in a grounded-grid circuit?

(bg) How do the input impedances of the grounded-cathode and cathode-follower stages compare?

(bh) How do the output impedances of the grounded-cathode and cathode-follower stages compare?

LOOKING AHEAD

In later lessons many of the circuits you have studied in this lesson will be discussed in much more detail. The purpose of this lesson is to give you a look at some of the circuits in which tubes are used in order to help you better understand how tubes work. Later you will see that there are many other important things to be considered in the various stages that we have studied.

It is important for you to understand how all the various types of circuits commonly found in electronic equipment operate. You will remember that we have pointed out several times that complex electronic equipment is made up of a large number of simple circuits connected to work together. If you understand how the individual circuits work, you will be able to understand how they work together. If you understand how a piece of electronic equipment is supposed to work, you should have no difficulty locating a defect in the equipment.

Answers To Self-Test Questions

(a) A Class A amplifier is an amplifier that is biased to operate midway between zero bias and cut-off bias.

(b) A Class B amplifier is an amplifier operated with the bias approximately at cut-off.

(c) A Class C amplifier is an amplifier operated with a bias two to four times cut-off bias.

(d) Voltage amplifiers and power amplifiers.

(e) A voltage amplifier is an amplifier that is designed to take a small signal voltage and amplify it to a larger signal voltage.

(f) Voltage amplifiers are Class A amplifiers.

(g) The chief disadvantage of a Class A power amplifier is its relatively poor efficiency.

(h) The maximum theoretical efficiency that can be obtained from a Class A power amplifier is 50%. However, usually efficiencies in the order of 30% to 35% are all that is practical to obtain.

(i) A single Class B audio amplifier cannot be used because it produces only one half of an input cycle. However, two tubes can be used as a Class B power amplifier providing they are arranged so that one tube reproduces one half of the input cycle, and the other tube reproduces the other half of the input cycle.

(j) Class C power amplifiers are used to amplify radio-frequency signals.

(k) A Class AB amplifier is an amplifier operated with a bias between the value that would normally be used for a Class A amplifier and the value that would normally be used for a Class B amplifier. Sometimes Class AB amplifiers are designated as Class AB_1 or Class AB_2. A Class AB_1 amplifier does not draw any grid current and hence is very much like a Class A amplifier. A Class AB_2 amplifier does draw grid current and hence is operated closer to a Class B amplifier.

(l) Voltage amplifiers and power amplifiers.

(m) See Fig. 6.

(n) The grid-bias battery can be eliminated in a voltage amplifier by inserting a suitable resistor in the cathode circuit of the tube. Current flowing from B- through the resistor, to the cathode of the tube, then from the cathode to the plate and back to the B+ will develop a voltage drop across the cathode resistor. If the cathode resistor is of the correct size, the cathode will be sufficiently positive with respect to ground to provide the required bias. The grid will be essentially at ground potential and therefore negative with respect to the cathode. If the voltage between the cathode and ground is equal to the required grid bias, the tube grid will have the correct negative bias applied to it.

(o) Electrons accidentally striking the grid of the tube flow through the grid resistor back to ground and to the cathode of the tube. If the grid resistor is made too large, these electrons will develop an undesired voltage across this resistor. The size of the resistor is limited by the number of electrons striking the grid of the tube.

(p) The capacitor is used to hold the voltage across the cathode-bias resistor constant. The signal applied to the grid of the tube causes an ac current, which is superimposed on the dc current flowing through the tube, to flow in the cathode circuit. This ac current in effect flows through the bypass capacitor so that the voltage across the bias resistor does not vary.

(q) The output transformer is a step-down transformer. The transformer is primarily an impedance-matching device; it matches the low-impedance speaker to the higher impedance of the tube.

(r) The complete statement is as follows: a power amplifier using one tube such as shown in Fig. 8 is called a single-ended stage, whereas one using two tubes such as shown in Fig. 9 is called a double-ended stage.

(s) A voltage amplifier. The radio frequency amplifier used in a radio or television receiver is used to build up or amplify the voltage of a weak signal picked up by the antenna.

(t) C3 is a screen bypass capacitor. It is used to place the screen at signal ground potential so that the screen will provide maximum shielding between the plate and grid of the tube.

(u) A Class C power amplifier has the best efficiency; a Class A amplifier has the poorest efficiency.

(v) Both detectors and rectifiers work on the principle of allowing current to flow through them in only one direction.

(w) Sixty pulses per second. The rectifier will conduct once each cycle and therefore on a 60-cycle power line you will get 60 pulses per second.

(x) 120 pulses per second. A full-wave rectifier conducts on each half cycle. Since there are two half cycles in each cycle, on a 60-cycle power line you will get 120 pulses.

(y) The increased number of pulses per second from a full-wave rectifier makes it easier to filter the pulsating dc to pure dc. Thus smaller filter chokes and filter capacitors may be used with a full-wave rectifier rather than with a half-wave rectifier.

(z) See Fig. 14.

(aa) The variation in the strength of the signal being received causes the amplitude of the pulses flowing through a diode detector to vary. This variation in signal received is the modulation or intelligence signal that is placed on the rf carrier.

(ab) In the grid circuit. In a grid-leak detector the grid and cathode act like a diode tube. The grid acts like a plate, and when the input signal swings the grid positive, current flows from the cathode to the grid of the tube. When the input signal swings the grid negative, no current flows from the cathode to the grid.

(ac) We can compare the grid-leak detector to a diode detector followed by a triode amplifier stage.

(ad) Part of the signal from the output of the oscillator is fed back to the input circuit. This signal is called the feedback signal or simply feedback.

(ae) See Fig. 17.

(af) The grid bias is developed across the grid resistor, R1. This bias is self-regulating; in other words, if the output of the stage tends to increase, the bias will increase to hold the output down. If the output tends to decrease, the bias will decrease to permit the output signal to increase.

(ag) C3 and L1.

(ah) Feedback is controlled by the feedback capacitors C3 and C4. These capacitors form a voltage divider network to control the amount of signal fed from the output back to the grid-cathode, which forms the input circuit.

(ai) L1, C5, C3 and C4 control the oscillator frequency.

(aj) A sine-wave signal.

(ak) The output signal from a multivibrator is essentially a square wave.

(al) C1 - R2 and C2 - R3 control the multivibrator frequency.

(am) The multivibrator shown in Fig. 19 is called a plate-coupled multivibrator. It is given this name because the energy necessary for oscillation is fed between the tubes from the plate of each tube to the grid of the other tube.

(an) The mixer stage. This stage is sometimes referred to as the mixer-oscillator stage because the two functions of mixing and oscillation are performed in a single tube in a five-tube superheterodyne receiver.

(ao) The oscillator frequency is higher than the frequency of the incoming signal.

(ap) 1956kc. The i-f frequency in a receiver is equal to the difference between the local oscillator frequency and the frequency of the incoming signal. If the oscillator operates higher than the incoming signal, then the oscillator frequency must be equal to the frequency of the incoming signal plus the frequency at which the i-f amplifier operates. Therefore, in this case the oscillator frequency must be 1500 + 456 = 1956 kc.

(aq) The second grid acts as the plate of the triode section of the tube which in turn is the oscillator section.

(ar) The primary winding along with the capacitor connected across it form a parallel-resonant circuit because the signal is fed to the coil and capacitor in parallel. The secondary circuit consisting of the coil and capacitor form a series-resonant circuit because the signal is induced in series with the turns of the coil.

(as) A remote cut-off tube is a tube with a grid specially constructed so that it takes a rather large negative bias to cut off the flow of plate current through the tube. This type of tube is normally used in i-f amplifier stages through which an automatic volume control voltage is fed.

(at) The audio signal voltage is developed by the second detector across a 220 pf capacitor C8.

(au) Bias for the first audio stage is developed across the 6.8 megohm resistor, R8. As you will remember, a few of the electrons leaving the cathode of the tube and travelling toward the plate will actually strike the grid of the tube. These few electrons flowing through R8, which has a high resistance, will develop a small voltage across this resistor having such a polarity that the grid end is negative. This voltage biases the stage.

(av) The power amplifier stage is a Class A power amplifier.

(aw) The chances are good that one of the tubes may have a burned-out heater. Since the tubes are connected in series, if any tube has a burned-out heater this will open the circuit and none of the tubes will light. However if all the tubes are good, there is a possibility that the on-off switch, the line cord, or the line-cord plug may be defective. In addition, the socket into which the receiver is plugged may be at fault. However, in most receivers of this type, when all the tubes fail to light, the heater of one of the tubes is burned out.

(ax) The signal source is placed between grid and cathode, and the load is placed in the plate circuit of the tube.

(ay) The output signal will be greater than the input signal, and the output signal will be 180° out of phase with the input signal.

(az) A cathode follower.

(ba) The signal source is placed between grid and ground, and the load is placed between the cathode and ground.

(bb) The output signal is always lower in amplitude than the input signal. We say that the gain of a cathode follower is less than one - in other words, there is a loss in signal amplitude in this stage.

(bc) In a cathode follower the output signal will be in phase with input signal.

(bd) Where it is important not to change the phase of the signal, and in impedance-matching applications.

(be) The input signal is placed between the cathode and ground, and the load is placed in the plate circuit.

(bf) In a grounded-grid amplifier the grid is grounded and acts to shield the output circuit from the input circuit and by so doing prevents feedback which could cause oscillation.

(bg) Input impedances of both types of stages are high.

(bh) The output impedance of a grounded-cathode stage is comparatively high, but the output impedance of a cathode-follower stage is low.

Lesson Questions

Be sure to number your Answer Sheet B111.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. Which element of a tube is operated at signal ground potential when the tube is used as a cathode follower?

2. Which class of amplifier is operated with a bias several times cut-off bias?

3. What will normally happen to the average plate current of a properly biased Class A amplifier when a symmetrical ac signal is applied to the amplifier?

4. What is the purpose of R2 in Fig. 6?

5. In the circuit shown in Fig. 11, what current develops bias for the stage?

6. Which type of rectifier, a half-wave or a full-wave, requires more filtering to get pure dc?

7. In which circuit does detection occur in the grid-leak detector?

8. On what basic principle do oscillators work?

9. What does avc found in modern receivers do?

10. What type of amplifier can be used to match a high impedance to a low impedance?

TOMORROW NEVER COMES

The fellow who coined the phrase, "Don't put off till tomorrow what you can do today," was one of the world's wisest men. As a sure-fire formula for success, this certainly hits the nail on the head.

Perhaps NRI men wonder why I repeat this warning so often. It is because I am convinced that of all the reasons for failure, this habit of "putting off" is the greatest. We can always find a good excuse for "putting off." We can easily convince ourselves that we are too tired (or lazy?), or that we don't feel well, or that it's too hot, or that we have too much to do. The reason we are "putting off" is not important; the fact that we are "putting off" is.

The best -- in fact, the only -- way to overcome this temptation is never to succumb to it. The first time you find yourself saying, "I'll skip studying tonight and do twice as much tomorrow," is the time to study twice as hard as usual. Soon you will no longer have to fight temptation; it will no longer exist. It won't occur to you to say, "I'll do it tomorrow." Instead, you will automatically say, "I'll do it now."

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-211

ACHIEVEMENT THROUGH ELECTRONICS

HOW TRANSISTORS WORK

B112

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

HOW TRANSISTORS WORK

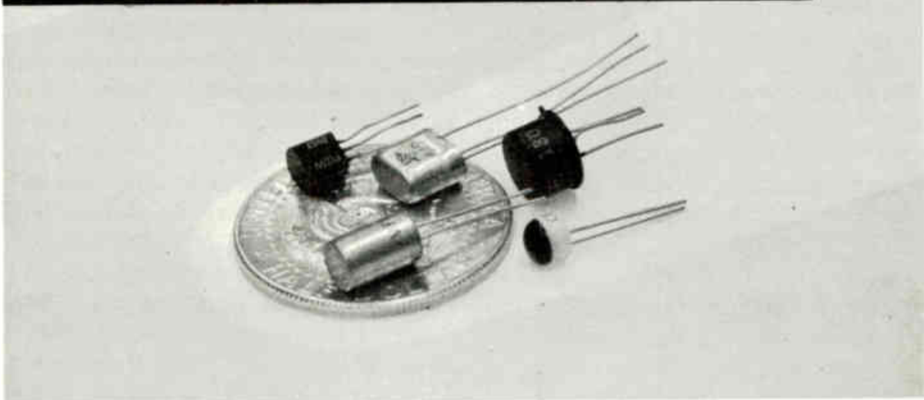
B112

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

PII3M375 **1975 EDITION** **Litho in U.S.A.**

HOW TRANSISTORS WORK



In the preceding lessons you studied tubes, and you saw how they are used in different circuits. In this lesson you will study semiconductor devices - these devices have already replaced tubes in many important applications and are rapidly moving into new areas that were once dominated entirely by tubes.

An example of the importance of semiconductor devices can be seen in entertainment-type equipment. Just a few years ago all the rectifiers used in this equipment were vacuum tubes. Today, however, the vacuum tube is no longer used for this purpose; rectifiers in the modern entertainment-type equipment are all semiconductor devices.

Semiconductor devices used as rectifiers have two elements and are called diodes just as two-element vacuum tubes are called diodes. Semiconductors used to amplify signals usually have three or more elements and are called transistors. There are a large number of different types of transistors available today, but for the most part these transistors can be classified into two types, the NPN transistor and the PNP transistor. If you understand how these two transistor types work, you should have little difficulty understanding how all others work and any new transistors that might be introduced in the future. You will run into many different types of transistors identified by different names, but these names usually refer to the method used in manufacturing the transistor rather than the manner in which it operates.

There are some similarities between tubes and semiconductors. A two-element vacuum tube can be used to change an alternating current to a direct current; a two-element semiconductor can be used for the same purpose. A triode vacuum tube can be used to amplify a signal; a transistor can be used for the same purpose. However, this is where the similarity ends. Most tubes are vacuum devices; in other words, all the air and gas have been evacuated from inside the tube. On the other hand, a semiconductor is a solid device and there is no space between the elements in it. We have a current flow through a vacuum in a tube, but we have a current flow through a solid in a semiconductor.

The importance of semiconductors cannot be overemphasized. They have completely supplanted the vacuum tube in portable radio receivers and in automobile receivers. Almost all high fidelity and stereo equipment manufactured today uses semiconductors exclusively - the only tube-operated equipment of this type you are likely to encounter is equipment that is several years old. Semiconductors are finding their way into television receivers and it is probably just a matter of time before they completely replace the vacuum tube.

Semiconductors have several advantages over the vacuum tube. Perhaps one of the most important advantages is that they do not require any heater or filament power. Not only is this a power saving in the operation of the equipment, but it also removes considerable heat from the equipment. Heat is probably the thing that causes the most damage to parts in electronic equipment. Thus with the removal of the heater or filament power from the equipment, other components such as capacitors, etc. will last longer.

Semiconductors are very rugged. They are solid devices and hence not subject to breakage from mechanical shock as tubes are. An important advantage of transistors is that they will operate on a comparatively low voltage, and this usually results in some reduction of the power required in the equipment.

Although semiconductors have many advantages over vacuum tubes, they do have some disadvantages. One disadvantage is that it is usually not possible to get as high a gain in an amplifier stage using a transistor as it is in a similar stage using a tube. Therefore to get the equivalent gain, more transistor stages are required than vacuum-tube stages. Another disadvantage of the transistor is that its characteristics are not as constant as those of a vacuum tube. In other words, you are more likely to run into difficulty replacing a transistor than you are in replacing a tube because the replacement transistor's characteristics might be considerably different from the characteristics of the original transistor. Another disadvantage of both diode semiconductors and transistors is that their characteristics can vary appreciably with changes in temperature. As a matter of fact, some semiconductor devices are easily destroyed by too much heat.

In spite of the fact that semiconductor devices have some disadvantages when compared to vacuum tubes, their advantages more than outweigh the disadvantages and their importance in the field of electronics is continually growing. Therefore it is important that the technician have a good understanding of semiconductor fundamentals, how they are used, and how they operate. Before going ahead to see how semiconductors are used as rectifiers and amplifiers, we need to know more about certain types of atoms, in order to understand how these devices work.

Semiconductor Fundamentals

You have already learned that certain materials will conduct electricity readily and that some materials will hardly pass any electric current at all. The materials that will conduct current readily are called conductors and those that will not conduct current are called insulators. Midway between the two types of materials is a group of materials called semiconductors. These materials are not good conductors, nor are they particularly good insulators. Two examples of semiconductor materials are germanium and silicon. These are the materials that we will be mostly concerned with in this section. Both diode semiconductors and transistors are made from germanium and silicon. A new material that shows promise for use in semiconductors is gallium arsenide. It's likely that this material will be used in semiconductors in the future. Before going ahead with our detailed study of semiconductor materials, let us review a few important facts about conductors and insulators.

CONDUCTORS AND INSULATORS

You will remember that all materials are made up of atoms. An atom is the smallest particle of a material that retains the characteristics of the material.

In the center of the atom is the nucleus. This nucleus contains a positive charge. The number of positive charges on the nucleus distinguishes one material from another. In other words, the nucleus of a copper atom does not have the same number of positive charges as the nucleus of an iron atom.

Each atom normally has enough electrons, which have a negative charge, to exactly neutralize the positive charge on the nucleus. Thus, the hydrogen atom which has a nucleus with one positive charge will have one electron, and the helium atom which has a positive charge of two in the nucleus will have two electrons. Another atom that has a nucleus with 30 positive charges will have 30 electrons to exactly neutralize the positive charge on the nucleus.

The electrons in an atom arrange themselves in shells around the nucleus. The total number of electrons will normally be just enough to neutralize the charge on the nucleus. However, there is a maximum number of electrons that can be forced into each shell. In the first shell around the nucleus, the maximum number of electrons is 2. In the second shell, the maximum number of electrons is 8, and in the next shell the maximum number of electrons is 18. A shell can have less than the maximum number of electrons, but not more than the maximum number.

Conductors.

An example of an atom in a conductor is shown in Fig. 1. We have drawn the shells in the form of rings, but remember that this atom actually has three dimensions, not two. Notice that in this atom there are two electrons in the first shell,

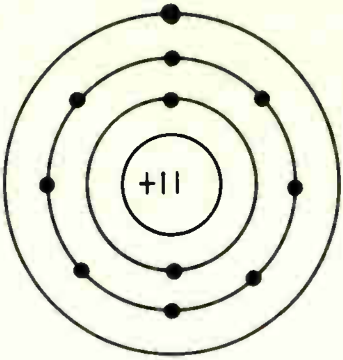


Fig. 1. An atom of a conductor.

8 electrons in the next shell and only one electron in the third shell. The outer shell is called a valence shell. The single electron in the third shell, which is called the valence electron, is not very closely bound to the nucleus; it can easily be removed from the atom. Thus a material of this type has a large number of electrons that can easily be removed from their atoms. When these electrons are forced to move in one direction we have a current flow. Thus a material that has only one or two electrons in an outer shell that could have many more, is a conductor, because the one or two electrons in the outer shell are not closely bound to the nucleus.

Insulators.

An atom of an insulator is shown in Fig. 2. Notice that in this atom

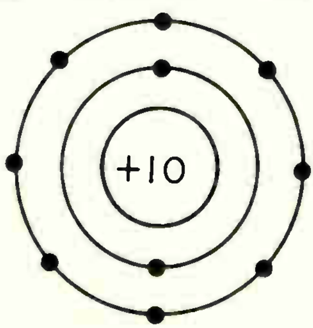


Fig. 2. An atom of an insulator.

there are two electrons in the first shell, and 8 electrons in the second shell. Both the shells are completely filled and will be closely bound to the nucleus. This means that it is very difficult to get one of these electrons out of an atom and therefore this material is an insulator or nonconductor.

Remember the important difference between conductors and insulators. A conductor is a material that has one or two electrons in the outer shell that are not closely bound to the nucleus, whereas an insulator is a material in which the outer shell of each atom is filled or almost filled so that the electrons are closely bound to the atom and cannot be easily removed. Because these electrons cannot be removed from the atom, this type of material normally will not conduct current, and hence is called an insulator.

SEMICONDUCTOR MATERIAL

A material that is classified as a semiconductor has electrical characteristics midway between those of a conductor and those of an insulator. The electrons in a semiconductor can be removed from their atoms when some type of external energy, such as voltage, heat, or light is applied to the material. Then the material acts like a conductor.

The most important semiconductor materials used for transistors are germanium and silicon. The first low-cost transistors were germanium transistors, but recent developments have lowered the cost of silicon transistors so that most of the new transistor types being introduced are silicon. Both germanium and silicon are very abundant elements, but neither is found in the pure state, and it is quite difficult to process them to the high state of purity required for use in transistors. The first transistors were made of germanium because techniques for getting pure germanium were developed first. However, now it is possible to refine silicon to the high degree of purity required, at a reasonable cost, and since silicon has several advantages over germanium for use in semiconductor devices it has in many ways replaced germanium.

In general, there is not too much difference between the operation of semiconductor devices made from germanium and those made from silicon. We will cover the important points of these devices made from both materials since you will run into semiconductor devices of both types.

The Germanium Atom.

The arrangement of electrons about the nucleus in a germanium atom is shown in Fig. 3A. The nucleus of the germanium atom has a positive charge of 32. Thus as you might expect, there will be 32 electrons revolving in the shells about the nucleus. There are two electrons in the first shell, eight in the second, eighteen in the third and four in the fourth shell. Thus, the first, second and third shells are filled, but there are only four electrons in the outer shell. However, these four electrons, which are called the valence electrons, are bound to the nucleus much more so than the one or two electrons found in the outer shell of a conductor.

The important electrons in the germanium atom, insofar as its use in semiconductors is concerned, are

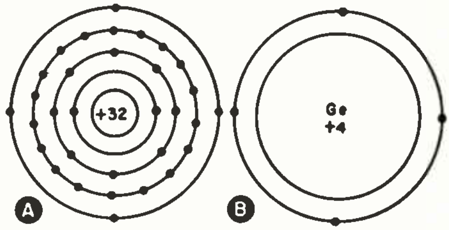


Fig. 3. (A) is the germanium atom with a charge of 32. (B) shows the simplified symbol.

the four electrons in the outer shell because the shell is not filled. The other electrons are bound so closely to the nucleus that they cannot easily be removed. Therefore, germanium is often represented as shown in Fig. 3B.

The Silicon Atom.

The arrangement of electrons about the nucleus in a silicon atom is shown in Fig. 4A. The nucleus of the silicon atom has a positive charge of 14. Therefore, there will be fourteen electrons revolving about the nucleus. There are two electrons in the first ring, eight in the second and four in the third. Thus the first and second rings are filled, but there are only four electrons in the outer shell. These four

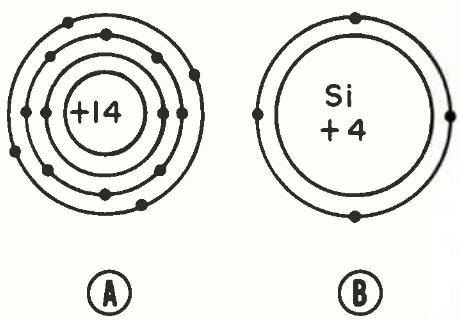


Fig. 4. (A) shows the silicon atom with a charge of 14. (B) shows the outer shell with four electrons.

electrons are the valence electrons like the four in the germanium atom and are bound fairly closely to the nucleus. As in the case of the germanium atom, the four electrons in the outer ring are the ones that are of importance in the use of silicon in semiconductors.

Notice the similarity between the silicon and germanium atoms. In both atoms, the outer shell or ring has four electrons, and all the other shells are filled.

The tendency of some materials like silicon and germanium, that do not have the outer shell completely filled with electrons, is to get additional electrons to fill up the outer shell. In pure germanium and silicon, the electrons in the outer shell of one atom are bound as closely to that atom as the four electrons in the outer shell of another atom. Therefore one atom cannot pull electrons away from another atom. Instead, two nearby atoms will share one outer electron from each atom. In other words, two atoms of germanium may share electrons as shown in Fig. 5A; and two atoms of silicon may share electrons as shown in Fig. 5B. By sharing electrons in this way, each atom will partly fill its outer shell. This pair of shared electrons, one from each of two atoms, is called "a covalent bond".

In order to try to fill its outer ring with electrons, a single germanium atom or a single silicon atom will establish covalent bonds with four other atoms. This arrangement of atoms in a piece of germanium is shown in Fig. 6A. A similar arrangement of atoms in a piece of silicon is shown in Fig. 6B. These pieces of silicon and germanium are called crystals and the way in which they are arranged is called a lattice structure. Each atom shares each of its four valence electrons with one valence electron of another atom to form these bonds.

INTRINSIC CONDUCTION

Even at comparatively low temperatures, there is heat energy in all materials. This energy is sufficient to cause a few of the electrons to move out of their proper place in the lattice structure of either the germanium crystal or the silicon crystal and become free electrons. These free electrons are available for conduction of electric current. The number of free electrons available is much higher in germanium than it is in silicon.

When one of these electrons moves out of its position in the lattice structure, it leaves an empty space in the crystal lattice. This empty space is called a hole. An electron from a

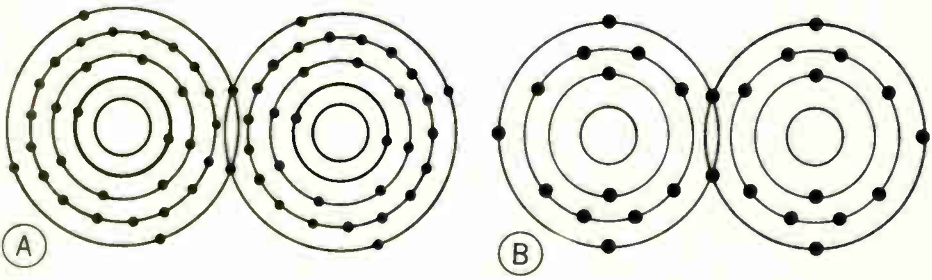
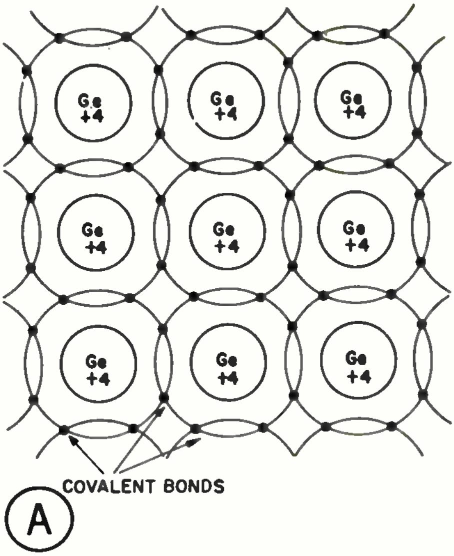


Fig. 5. The sharing of two electrons by two germanium atoms is shown at A; by two silicon atoms at B.



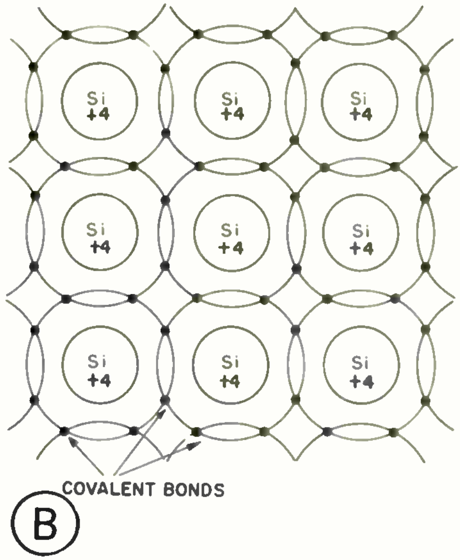


Fig. 6. The lattice structure of germanium is shown at A; of silicon at B.

nearby atom can move into this hole thus creating a new hole at the place it left. Another electron may move out of still another atom to fill this new hole, leaving behind it a hole. This movement of an electron to fill a hole thus creating a new hole in the place it left makes it look as if holes themselves move. Furthermore, since the hole represents a missing electron, it has a positive charge.

In a piece of germanium or silicon, the electrons are in a constant state of motion about their atoms. If in its movement an electron comes closer to a hole than to its own atomic nucleus, it will be strongly attracted to the hole and will leave its atom. When there is no voltage applied across the crystal, the movement of a hole or an electron is a random movement. Holes and electrons may move in any direction.

If heat or some other form of external energy is applied to the crystal, the resistance of the material is reduced. This happens because more electrons are freed by the energy applied to the crystal. In addition, the speed of the random movement is increased.

The movement of an electron out of an atom forms a hole in the atom. Thus, whenever an electron is freed from an atom a hole is formed. This free electron and the hole it forms are called a "hole-electron pair". The formation of hole-electron pairs is a continuous process. Also the filling of holes by electrons is a continuous process. In other words, the process of an electron leaving its atom and forming a hole, and another electron moving in to fill the hole and in so doing creating a new hole, is a continuous process. The conduction of electricity in pure germanium or pure silicon crystals due to the formation of hole-electron pairs is called the intrinsic conduction.

The conductivity of a germanium crystal or a silicon crystal, which is the ability of the material to conduct an electric current, depends on the average length of time an electron is free and on the number of free electrons. We mentioned previously that there are more electrons free if external energy, such as heat, is applied to the material. Therefore the conductivity rises as the temperature of the material is increased.

This type of conduction is much higher in germanium than it is in silicon. As an example, if we had a germanium crystal exactly one centimeter on each side and measured the resistance across two parallel surfaces, we would find the resistance to be approximately 60 ohms. The resistance of an equivalent piece of silicon would be approximately 60,000 ohms. Thus intrinsic conduction is much higher in germanium than it is in silicon.

Intrinsic conduction in transistors is undesirable. It is kept as low as possible by holding the operating temperature of the material down. Transistors are also shielded from light because light is a form of energy and light striking the crystal will increase the intrinsic conduction. Since silicon has a much lower intrinsic conduction than germanium, semiconductors made from silicon are less affected by heat than are semiconductors made from germanium. This is one of the chief advantages of silicon over germanium as a semiconductor material.

In their pure forms, neither germanium nor silicon are useful in semiconductor devices. In fact, in spite of intrinsic conduction, neither material is a good conductor at room temperature; they are both fairly good insulators. To use these materials in semiconductors, controlled amounts of other selected elements called impurities are added to the crystals to alter their characteristics. By adding these materials we can produce two types of silicon and two types of germanium. They are called N-type and P-type. Now, let us study the characteristics of these two types of materials.

N-TYPE MATERIAL

N-type silicon or germanium can be produced by adding as an impurity an element that has five electrons in its outer ring. An example of this type of material is arsenic. Arsenic has a positive charge of 33 on the nucleus and has 33 electrons in the shells surrounding the nucleus. There are two electrons in the first shell, eight in the second, eighteen in the third and five in the fourth or outer shell. In other words, arsenic is just like germanium except that the nucleus has one more positive charge and there is one additional electron in the outer shell.

If a small amount of arsenic is added to the germanium, the arsenic atoms will form covalent bonds with the germanium atoms as shown in Fig. 7. However, to form the covalent bonds with its neighboring germanium atoms, the arsenic atom needs only four of the electrons in its outer shell. Therefore there will be one electron left over when the arsenic atom forms covalent bonds with the four neighboring germanium atoms. This electron is free to move about within the crystal in exactly the same manner as a single valence electron in a good metal conductor. The addition of arsenic, which produces these free electrons, greatly reduces the resistance of the material.

When a small amount of arsenic

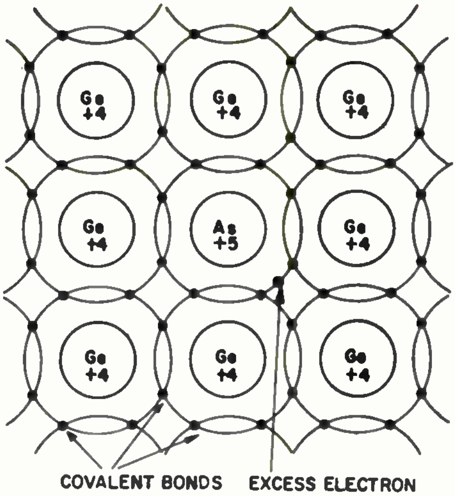


Fig. 7. Germanium with arsenic added.

is added to a silicon crystal exactly the same thing happens. The arsenic atom forms covalent bonds with the silicon atoms. As in the case of the germanium atom, only four of the electrons in the outer shell of the arsenic atom are used in forming these covalent bonds so there will be one electron left over.

When germanium or silicon have had an impurity added to them we say they have been doped. When semiconductor material has been doped with a material such as arsenic that results in there being excess electrons, we call it an N-type material. The N refers to the negative carriers, which are the free electrons. Arsenic is called a donor impurity because it donates an easily freed electron.

In addition to arsenic, other materials have been used as donors. Phosphorus, which has a total of fifteen electrons, can be used. The phosphorus atom has two electrons in the first shell, eight in the second and five in the third. Four of the electrons in the valence shell or ring will form covalent bonds with germanium or silicon atoms leaving a fifth electron free. Antimony, which has 51 electrons, also has been used as a donor. Antimony has two electrons in the first shell, eight in the second, eighteen in the third, eighteen in the fourth and five in the fifth or valence shell.

P-TYPE MATERIAL

If instead of adding a material with five electrons in its valence shell, we add a material with only three electrons in the valence shell, we have a situation where the impurity added to the silicon or germanium has one less electron than it needs to establish covalent bonds with four neighboring atoms. Thus, in one covalent bond there will be only one electron instead of two. This will leave a hole in that covalent bond.

A material that is frequently used for this purpose is indium. Indium has 49 electrons, two in the first shell, eight in the second, eighteen in the third, eighteen in the fourth and three in the fifth or valence shell.

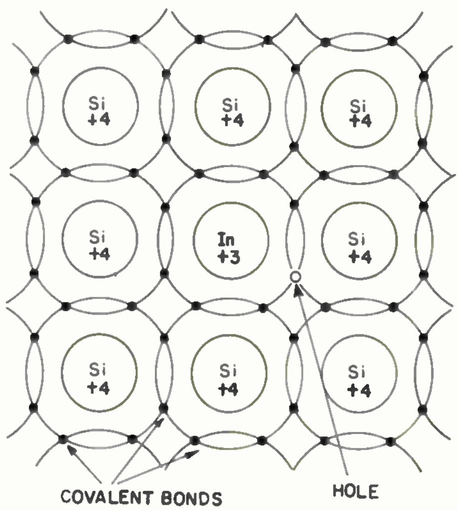


Fig. 8. Silicon with indium added.

The manner in which indium forms covalent bonds with neighboring silicon atoms is shown in Fig. 8. It forms covalent bonds with germanium atoms in the same way.

We mentioned previously that even at comparatively cold temperatures there is some heat energy within the crystal and thus there will be a few free electrons moving about the crystal. These free electrons are strongly attracted to the holes in the covalent bond produced where an indium atom has replaced a silicon or a germanium atom. Thus an electron will move into a hole in the covalent bond producing a new hole in another atom and giving the effect that the hole is moving as shown in Fig. 9.

Since a hole in the crystal actually represents a shortage of an electron, it is an area with a positive charge. Therefore when a semiconductor material has been doped with a material such as indium that produces holes in the lattice structure, we call it a P-type material.

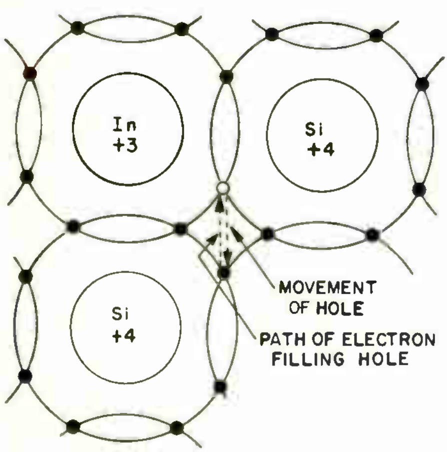


Fig. 9. When an electron fills a hole, another hole will apparently move to where the electron was.

P stands for positive; since holes represent a shortage of an electron we say they act as positive carriers. The indium is called an accepter impurity because its atoms leave holes in the crystal structure that are free to accept electrons. In addition to indium, boron and aluminum are also used as accepter impurities. Boron has 5 electrons, two in the first shell and three in the second which is the valence shell. Aluminum has 13 electrons, two in the first shell, eight in the second and three in the third or valence shell.

CHARGES IN N-TYPE AND P-TYPE MATERIAL

When a donor material such as arsenic is added to germanium or silicon, the fifth electron in the valence ring of the arsenic atom does not become part of a covalent bond. This extra electron may move away from the arsenic atom to one of the nearby germanium or silicon atoms.

The arsenic atom has a charge of +33 on the nucleus and normally has 33 electrons to neutralize this charge. When the electron moves away from the atom there will be only 32 electrons to neutralize the charge on the nucleus, and as a result there will be a small region of positive charge around the arsenic atom. Similarly, the excess electron that has moved into a nearby germanium or silicon atom will provide an excess electron in the atom. In the case of the germanium atom there will be a total of 33 electrons around a nucleus requiring only 32 electrons to completely neutralize it, and in the case of the silicon atom there will be 15 electrons around the nucleus requiring only 14 electrons to neutralize it. This means that the atom will have an extra electron so that there will be a region of negative charge around this atom.

It is important to notice that although there is a region of positive charge around the arsenic atom after the electron has moved away, and a region of negative charge around the germanium or silicon atom taking up the extra electron, the total charge on the crystal remains the same. In other words, a given crystal will have a net charge of zero. This means that there will be exactly enough electrons to neutralize the positive charges on the nuclei on the various atoms. But because some of the electrons may move about in the crystal, there will be regions in the crystal where there are negative charges and other regions where there are positive charges, even though the net charge on the crystal is zero.

In a P-type material to which material such as indium has been added we will have a similar situation. You will remember that the indium atom has only three electrons in its valence ring. These are all that were needed to neutralize the positive charge on the nucleus. However, with only three electrons in the valence ring, there is a hole in one of the covalent bonds formed between the indium atom and the four adjacent germanium or silicon atoms. If an electron moves in to fill this hole, then there is one more electron in the indium atom than is needed to neutralize the charge on the nucleus. Thus there will be a region of negative charge around the indium atom. Similarly, if one of the germanium or silicon atoms has given up an electron to fill the hole in the covalent bond, then the atom which has given up the electron will be short an electron so that there will be a region of positive charge around this atom. Again, while this giving up of an electron by a germanium atom and the acceptance of an electron by the indium atom ionizes or charges both atoms involved, the net charge on the crystal is still zero. We simply have one atom that is short an electron and another atom that has one too many. The crystal itself does not take on any charge.

These ionized atoms produced in both the N-type and the P-type germanium and the N-type and the P-type silicon are not concentrated in any one part of the crystal, but instead are spread uniformly about the crystal. If any region within the crystal were to have a very large number of positively charged atoms, these atoms would attract free electrons from other parts of the crystal to neutralize part of the charged atom, so that the charge would be spread uniformly about the crystal. Similarly, if a large number of atoms within a small region have had an excess of electrons, these electrons would repel each other and spread throughout the crystal.

Both holes and electrons are involved in conduction at all times. Holes are called positive carriers and electrons negative carriers. The one present in greatest quantity is called the majority carrier; the other is the minority carrier. In an N-type material, electrons are the majority carriers and holes are the minority carriers, whereas in a P-type material, holes are the majority carriers and electrons the minority carriers.

SUMMARY

This is a very important section of this lesson. You have covered many of the fundamentals of semiconductors on which we will build the remainder of the lesson. It is important that you understand the basic theory of semiconductors in order to be able to understand how semiconductor diodes and transistors work. We will summarize the important points that were covered in the preceding section.

If any of these points are not clear, you should go back and study the lesson again until they are clear. If you understand the first section of the lesson, you should be able to understand the material following without too much difficulty. However, if you do not understand what has been covered previously, you will have difficulty understanding what is to follow.

Pure semiconductor material such as germanium or silicon is a very poor conductor. In fact, it is an insulator if it is protected from all outside sources of energy. However, even at room temperature there is enough heat present in germanium and silicon to produce some electron and hole movement. The movement is much greater in germanium than it is in silicon.

An electron movement out of a covalent bond in a germanium or silicon atom leaves a hole in that bond. The hole will attract an electron from a nearby atom, producing a hole in that atom. Thus, both the hole and the electron appear to move. The holes are positive carriers and the electrons negative carriers of electricity. This formation of hole-electron pairs is undesirable in transistors and steps are taken to keep it as low as possible. The formation of hole-electron pairs increases as the temperature increases and is a much more serious problem in germanium-type semiconductor material than in silicon-type semiconductor material.

Semiconductor materials can be doped by adding small amounts of impurities. If a material with five electrons in the valence ring is added, the material is called a donor-type impurity. This type of material has one electron left over after it forms covalent bonds with four neighboring germanium or silicon atoms. Thus there will be an excess of electrons. We then refer to this kind of material as an N-type material.

If the germanium or silicon is doped with an impurity, called an accepter impurity, having three electrons in the valence shell, the impurity forms covalent bonds with the four neighboring germanium or silicon atoms. However, there will be a hole in one of the covalent bonds because the impurity has only three electrons available to form covalent bonds with four neighboring germanium and silicon atoms. This type of germanium or silicon is called P-type because there will be holes in the material, and these holes act as positive carriers.

Another point to remember is that when an electron is freed or when a hole captures an electron, the atoms involved become charged, or ionized. Thus throughout both N-type and P-type germanium or N-type and P-type silicon we have small regions of charge. However, the net charge on the crystal is zero and the charged regions are evenly distributed throughout the material.

SELF-TEST QUESTIONS

(a) How many electrons are there in a valence shell or ring in a silicon atom?

(b) What are the two types of material most widely used in semiconductor devices?

(c) What is meant by a covalent bond?

(d) How many covalent bonds will a single germanium or silicon atom establish?

(e) What is intrinsic conduction?

(f) Is intrinsic conduction desirable?

(g) In which type of conductor material, germanium or silicon, is intrinsic conduction the greatest?

(h) What is the greatest cause of an increase in intrinsic conduction in germanium?

(i) Which semiconductor material, silicon or germanium, has greater resistance?

(j) What is an N-type material?

(k) What is a donor material?

(l) Name two materials used as donors.

(m) What is P-type material?

(n) What is an accepter impurity?

(o) Name three types of accepter material.

(p) In N-type material what is the majority carrier?

(q) What are the majority carriers in the P-type material?

(r) When a donor impurity such as arsenic loses an electron in a semiconductor material, what happens to the arsenic and to a nearby atom that gains the electron insofar as their relative charge is concerned?

(s) Although there are small areas that have positive and negative charges in a doped semiconductor, what is the overall charge on the crystal?

Current Flow in Semiconductors

In order to understand how transistors operate, there are several new ideas that you must master. First, you must understand how current flows through both N-type and P-type semiconductor materials. Current flow through an N-type material is not too different from current flow through metals, which you have already studied. However, there is quite a difference in the way current flows through a P-type material.

When a P-type material is placed next to an N-type material, we have what is called a junction. The action that occurs at the point of contact between these two different types of materials is extremely important. It is this action that makes the transistor possible.

In this section of the lesson we will study how current flows through N-type and P-type materials. We need to understand current flow through both types of germanium or silicon to be able to understand how a junction works. In a later section we will see how a junction works. Later, we will see what happens in a transistor, which has two junctions.

This section is extremely important, and you should be sure that you understand it completely. Once you understand this material, it will be a simple step to see how transistors can be used to amplify signals.

DIFFUSION

As we have mentioned, adding impurities to pure germanium or silicon adds free electrons or holes. You might at first think that when there is no voltage applied there would be no motion of the free holes and electrons. However, this is not true--as you learned when we discussed intrinsic conduction, there is a certain amount of energy present in the crystal. This energy might be due to the temperature of the crystal, because as we pointed out before, even at room temperature the crystal does have heat energy. Motion of the free holes or electrons due to energy of this type is at random; in other words there is no net movement in any one direction. Holes move one atom at a time, and any hole may move from its starting location to any of the surrounding atoms. This means that a hole may start off in one direction as it moves from one atom to another, and then may move in almost the opposite direction as it moves to still a third atom. Similarly, electron movement is in a random direction; a given electron may move in first one direction and then in another.

When electrons and holes are in motion, the different carriers are moving in different directions. Remember that when there is a hole in one atom, and an electron moves from another atom to fill that hole, a new hole appears in the second atom. In other words, the electron has moved from the second atom to the first, whereas the hole has moved in the opposite direction from the first atom over to the second atom that gave up the electron. The result is that the effective current flow of any one carrier is cancelled by the movement of the other carrier and the resulting current flow in any direction is zero.

This random motion of carriers is called diffusion. It goes on at all times in a crystal whether there is a voltage applied to the crystal or not. Every effort is made in the design of transistors to keep this diffusion as low as possible.

DRIFT

Another type of carrier movement in semiconductors is known as "drift". This is the type of movement that is obtained when a voltage is applied across the crystal. Since the manner in which current flows through N-type and P-type material is different, let's consider them separately.

N-Type Material.

In Fig. 10 we have shown an N-type crystal with a voltage applied to it. The voltage difference supplied by the battery provides a force which makes it easier for the electrons to move in one direction than in the other. In an N-type material, the electrons will be attracted by the positive terminal of the battery. Because in the N-type material the electrons greatly outnumber the holes, they will carry the current.

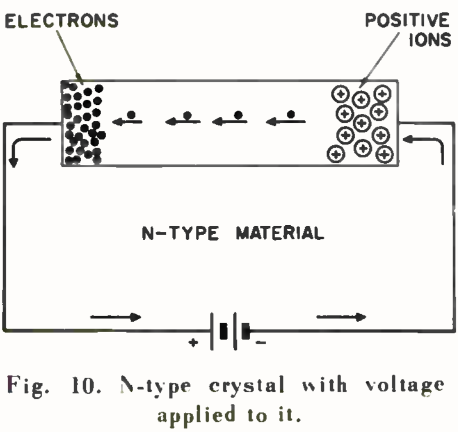


Fig. 10. N-type crystal with voltage applied to it.

When the electrons are attracted by the voltage applied to the positive terminal, they will move towards the positive terminal. When an electron moves away from the covalent bond that produced this free electron, it will leave behind an atom with a positive charge, which we call a positive ion. The electrons moving towards one end of the crystal set up a region that has a local negative charge, as shown in Fig. 10. This negative charge sets up a potential difference between that part of the crystal and the positive terminal of the battery. In other words the attraction of the positive battery terminal causes electrons to bunch up near the end of the crystal connected to the positive terminal. The electrons are drawn from the crystal into the wire connecting the crystal to the positive terminal of the battery by this potential difference.

Meanwhile, the electrons that have left the atoms at the other end of the crystal have left behind positive ions. This sets up a region of positive charge around the end of the crystal connected to the negative terminal of the battery so there will be a potential difference between the negative terminal of the battery and this region of positive charge. This potential difference will pull electrons from the wire into the crystal. These electrons replace the free electrons that were attracted to the positive terminal of the battery.

The number of electrons leaving the crystal at the end connected to the positive terminal of the battery will be exactly equal to the number of electrons entering the crystal at the end connected to the negative terminal of the battery. Since the crystal was electrically neutral be-

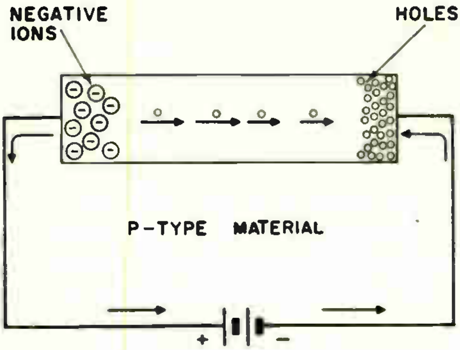


Fig. 11. P-type crystal with voltage applied to it.

fore the battery was connected and the number of electrons in it are constant, the crystal remains electrically neutral.

P-Type Material.

Conduction through P-type material is quite different from conduction through N-type material. In the P-type semiconductor, nearly all of the current is carried by holes. When a battery is applied to a P-type semiconductor, as shown in Fig. 11, the voltage causes the holes to drift towards the negative terminal. They are repelled by the positive potential applied to the one end of the material and attracted by the negative potential applied to the other end. When a hole starts moving away from the end of the material connected to the positive terminal of the battery, it moves because it is filled by an electron attracted from a nearby germanium atom.

When the hole in an accepter type atom is filled with an electron, the atom actually has one electron more than it needs to neutralize the charge on the nucleus. Thus, the atom has a negative charge, or in other words it becomes a negative ion. Negative ions that are formed near the end of the semiconductor that is connected to the positive terminal of the battery build up a region of negative charge at this end of the material. The extra electrons are drawn from these ions by the positive terminal of the battery, and a new hole is formed. These holes then drift towards the end of the semiconductor that is connected to the negative terminal of the battery, and build up a positive charge at this end of the semiconductor. This positive charge attracts free electrons from the external circuit. As a hole is filled with an electron, it disappears.

Thus in the P-type material, we have an electron flow in the external circuit from the negative terminal of the battery to the semiconductor, and from the semiconductor to the positive terminal of the battery. However, in the semiconductor itself, current flow is by means of holes, which drift from the end of the semiconductor that is connected to the positive terminal to the end that is connected to the negative terminal of the battery. Keep this point in mind, that even in the P-type material where conduction within the material is by holes (which are positive carriers) the current flow in the external circuit is by means of electrons and is in the conventional direction from the negative terminal towards the positive terminal of the battery in the external circuit.

There are several important differences between conduction in N-type semiconductors and conduction in P-type semiconductors. In both cases electrons flow from the external circuit into the crystal and then out of the crystal into the external circuit. However, in the N-type crystals, the excess electron produced when a donor atom forms covalent bonds with four germanium atoms is a free electron that can move about in the crystal. However, in the P-type material, the electrons are not free, but can move only to holes. Since a hole can capture an electron from any of its surrounding atoms, it is the hole that is free to move in any direction.

Another important difference between the N-type and the P-type materials is that a free electron moves approximately twice as fast as a hole. This affects the conductivity of the two types of semiconductor material. If we have two crystals, one an N-type and the other a P-type, if the N-type material has the same number of free electrons as the P-type has holes, the N-type will have a lower resistance because the free electrons can move approximately twice as fast as the holes in the P-type material.

SUMMARY

The important thing to remember from this section is that there are two types of carrier movement in semiconductors. The first is called diffusion and is simply a random movement of the carriers in the semiconductor material. The current flow produced by one carrier is cancelled by the movement of the other, and the resultant current flow in any direction is zero. Diffusion is the random motion of electrons or holes in a doped semiconductor due to the energy of the material.

The other type of movement we discussed is called drift. This type of conduction is produced when a potential is connected across a semiconductor. This potential can cause either electrons or holes to move within the semiconductor. In an N-type semiconductor, current flows through the semiconductor because of the movement of the free electrons produced by the donor atoms that have been added to the semiconductor material. In the P-type semiconductor, current flow through the crystal is by means of holes which are produced when an accepter-type impurity is added to the crystal.

In both cases current flow in the external circuit is from the negative terminal of the battery to the crystal and from the crystal to the positive terminal of the battery. In the N-type material, electron flow through the crystal is from the end connected to the negative terminal of the battery to the end connected to the positive terminal of the battery. In the P-type semiconductor, the holes flow from the end of the semiconductor connected to the positive terminal of the battery to the end of the semiconductor connected to the negative terminal of the battery.

The speed with which electrons move through N-type material is about twice the speed with which holes move through P-type material. Thus N-type material has better conductivity than P-type material, which means that N-type germanium will have a lower resistance than P-type germanium.

SELF-TEST QUESTIONS

(t) When an accepter-type impurity is added to a silicon or a germanium crystal, what type of carrier is produced in the crystal?

(u) What is diffusion?

(v) What is the name given to the movement of carriers in a semiconductor material when a voltage is applied across the material?

(w) What are the majority carriers in an N-type material and in what direction do they move when a voltage is applied across the material?

(x) What are the majority carriers in a P-type material and in what direction do they move through the material when a potential is applied across the material?

(y) When current is flowing through a crystal, will the crystal be charged?

(z) Is the rate of travel of electrons through N-type material the same as the rate of travel of holes through P-type material?

(aa) If you had two identical pieces of silicon and one was doped so that it was N-type material and the other doped so that it was P-type material, which would have the lower resistance?

Semiconductor Diodes

Just as there are diode tubes, there are also diode semiconductors. Some diode semiconductors are used as detectors: others are used as rectifiers in power supplies to change ac to pulsating dc. Diodes used as detectors are often referred to as signal diodes. Both germanium and silicon signal diodes are widely used. Diodes used for power rectification are almost exclusively silicon diodes. Relatively small silicon diodes can often handle considerably more current than a large rectifier tube.

A semiconductor diode is made by taking a single crystal and adding a donor impurity to one region and an accepter impurity to the other. This will give us a single crystal with a P section and an N section. Where the two sections meet, we have what is called a junction. Contacts are fastened to the two ends of the crystal so that a simple PN junction diode like the one shown in Fig. 12 is formed. For simplicity in the diagram we have represented the crystal as a box-like structure with one half being P-type material and the other half N-type material with a junction between the two sections.

This type of diode is called a junction diode. The action that takes place at the junction of the P-type crystal and the N-type crystal is what we will be most concerned with now. In order to understand how a junction diode works, you must learn something about the movement of electrons and holes near the junction. The movement of holes and

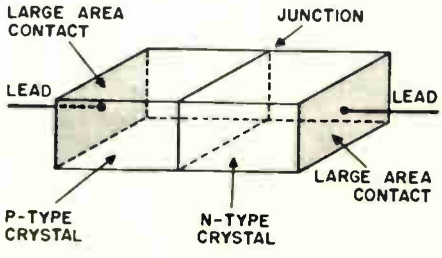


Fig. 12. Simple PN junction.

electrons will form what is called a depletion layer at the junction. Now let us see what the depletion layer is and how it is formed.

DEPLETION LAYER

Remember that in an N-type crystal there are free electrons, and in a P-type crystal there are free holes. Also remember that the electrons and holes are moving about the crystal with a random motion, called diffusion. In the PN junction diode, holes will be moving about in the P section and electrons in the N section. Some of the holes will cross over the junction from the P section into the N section and be filled by a free electron. Similarly, some of the electrons in the N-type material will diffuse across the junction and fill a hole in the P section.

When an atom in the N section loses an electron the atom becomes charged or ionized. It will have a positive charge because it will have one less electron than is needed to completely neutralize the charge on the nucleus. Thus electrons diffusing across the junction to fill a hole on the P side of the junction will leave behind atoms with a positive charge. At the same time, when an electron fills a hole on the P side, the atom will have one more electron than it needs to completely neutralize the charge on its nucleus, and therefore that atom will have a negative charge. Similarly, holes diffusing from the P side of the junction over into the N side will leave behind atoms with a negative charge. When the hole moves over to the N side, it will mean the atom into which it moves will have an electron missing and therefore it will assume a positive charge. When the hole leaves the P side of the junction because it has been filled by an electron, the atom that gains the extra electron will have a negative charge.

As a result of this diffusion across the junction, a region will build up around the junction called the depletion area. On the P side of the junction there will be an area where the holes are missing. On the N side of the junction there will be an area where electrons are missing; thus we get the name depletion layer.

The missing holes on the P side of the junction will result in a negative charge on the P side and the missing electrons on the N side will produce a positive charge on the N side of the junction. The negative charge on the P side of the junction will build up until it has sufficient amplitude to prevent any further electrons from the N side from crossing the junction to the P side. Remember that the negative charge built up on the P side of the junction will repel electrons from the negative side. Similarly, the positive charge built up on the N side of the junction will prevent holes from the P side from crossing the junction into the N-type material. Thus this area, which is called the depletion layer because it is short holes on one side and electrons on the other side, is also sometimes called the barrier layer, because the charges built up form barriers to prevent any further diffusion of holes or electrons across the junction. It is also sometimes called a potential barrier because a negative potential is built up on the P side of the junction and a positive potential is built up on the N side of the junction.

The action taking place at the junction is quite important and is illus-

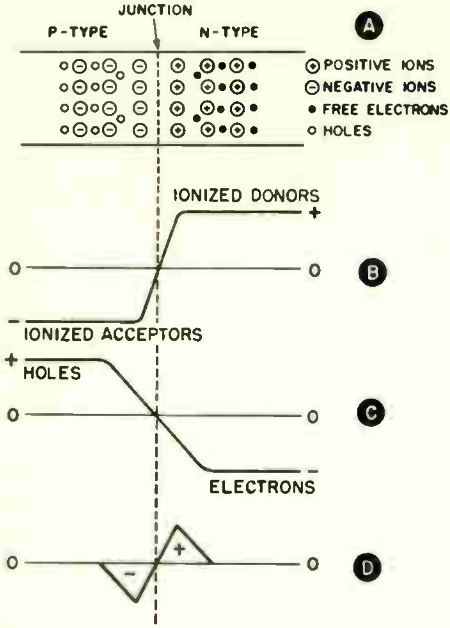


Fig. 13. (A) locations of ions and carriers at a PN junction; (B) charges at junction due to ionized impurity atoms; (C) carrier charges available; (D) resultant charges.

trated in Fig. 13A. On the P side of the junction we have shown ionized atoms that have a negative charge because the holes in these atoms have been filled by electrons. The holes have escaped and travelled or diffused across the junction into the N-type material. On the N side of the junction we have shown atoms that are ionized and have a positive charge. These atoms have a positive charge because they have lost electrons. These electrons have diffused across the junction into the P-type material. Thus we have a charged area at the junction. The negative charge on the P side of the junction prevents any further movement of electrons from the N-type material across the junction into the P-type material, and the positive charge on the N side of the junction prevents any further movement of holes from the P-type material across the junction into the N-type material.

The charge on the ions is shown in Fig. 13B. Notice that on the P side of the junction the atoms that have lost holes by gaining electrons have a negative charge. At the junction the potential drops to zero and then reverses on the N side where the ionized atoms have a positive charge because they have lost electrons.

In Fig. 13C we see the carrier charges which are available to neutralize the ionized atoms. At some distance from the junction there are holes with a positive charge. However, as we approach the junction, the concentration of these holes decreases because they are repelled away from the junction by the positive ions on the N side of the junction. On the N side of the junction at some distance from the junction we have many electrons available, but as we approach the junction, the charge drops to zero because these electrons are repelled away from the junction by the negative ions on the P side of the junction.

The resultant charges on the crystal are shown in Fig. 13D. As before, the crystal will have a tendency to remain neutral, or in other words not to have any charge. Some distance from the junction the atoms will have exactly the correct number of holes and electrons so that the net charge on the atoms is zero. As we approach the junction, the negative ions on the P side will result in an area in the crystal that has a negative charge. As we move closer to the junction, the charge will drop to zero so that at the junction itself the net charge on the atoms is zero. Then the charge builds up in a positive direction on the N side of the junction due to the ionized atoms that have lost electrons. As we move away from the junction we again reach a region where the atoms have exactly the correct number of electrons to neutralize the charges on the nucleus so the net charge in that area will be zero.

So far we have been discussing only the action of the majority carriers at the PN junction. However, there is one other important point we must consider in order to completely understand what happens at the junction. You will remember some time ago that we mentioned that holes and electrons are in a continuous state of motion in the crystal due to the energy of the crystal. For example, even at room temperature, the crystal contains a certain amount of heat energy and this energy is sufficient to cause motion of both electrons and holes. In the N-type material an electron will leave an atom creating a hole. This hole will be filled by an electron from another atom. Thus we have the continual formation of hole-electron pairs. Away from the junction, this formation of hole-electron pairs does not have any effect on the carrier concentration in the crystal. In other words, the holes will remain the majority carriers in the P-type region, and the electrons will remain the majority carriers in the N-type side of the crystal.

However, as we mentioned previously, both holes and electrons are involved in conduction at all times. There are minority carriers in both regions - holes in the N region and electrons in the P region. The holes produced in the N region near the junction will be attracted by the negative ions on the P side of the depletion layer at the junction and pass across the junction. These holes will tend to neutralize the ions on the P side of the junction. Similarly, free electrons produced on the P side of the junction will pass across the junction, and neutralize positive ions on the N side of the junction. This is an example of intrinsic conduction, conduction due to the formation of hole-electron pairs, and as we mentioned, this type of conduction is undesirable.

Now let us consider what happens due to the minority carriers crossing the junction. Holes crossing the junction from the N-type material to the P-type material tend to neutralize the negative ions on the P side of the junction. Similarly, electrons traveling from the P side of the junction to the N side of the junction tend to neutralize the positive ions on the N side of the junction. This flow of minority carriers across the junction weakens the potential barrier in the region around the atoms they neutralize. When this happens, the majority carriers are able to cross the junction at the location of the neutral atom. This means that the holes from the P side will cross over to the N side, and electrons from the N side will cross over to the P side.

The result is that we have both holes and electrons crossing the junction in both directions. The hole that crosses from the N side to the P side due to intrinsic conduction permits a hole to cross from the P side to the N side by diffusion. Similarly, an electron that crosses the junction from the P side to the N side due to intrinsic conduction permits another electron to go from the

N side to the P side by diffusion. The result of the holes and electrons crossing the junction in both directions is that these movements cancel each other, and the charge on the atoms at the junction remains the same. This movement of holes and electrons in both directions contributes nothing towards the net charge or current flow through the junction. However, the flow across the junction will produce a certain amount of heating; it will in effect use up a percentage of the total capacity of the junction to pass current so that the net result is to reduce the amount of useful current the diode can pass.

BIASED JUNCTIONS

If a battery is connected to the ends of a PN junction diode, the battery potential will bias the junction. If we connect the battery so that its polarity aids the flow of current across the junction, we call it a "forward-biased junction", whereas if we connect the battery so that the polarity opposes the flow of current across the junction, we say that it is a "reverse-biased junction". In both cases there will be some current flow through the junction, but as you might expect, with forward bias the current flow will be higher.

In order to understand how transistors work, you must understand both conditions of bias. You will study each condition separately, because the action that occurs at the junction is quite different in the two cases. In the operation of transistors both types of bias are used, and therefore it is important that you understand what happens in each case.

Forward Bias.

When we connect a battery to a junction diode with the polarity such that it aids the movement of majority carriers across the junction, we say that the diode is forward biased. A forward-biased junction is shown in Fig. 14. Here the positive terminal of the battery is connected to the P-type section and the negative terminal of the battery is connected to the N-type section. Now let us consider what happens to the depletion layer at the junction of the P and N-type material when the battery voltage is applied.

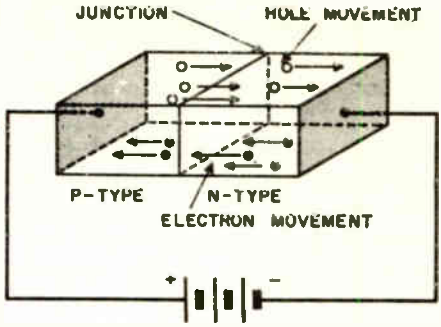


Fig. 14. Forward-biased junction.

The positive voltage connected to the end of the P-type crystal will repel holes towards the junction and attract electrons from the negative ions near it. The combination of holes moving towards the junction to neutralize charged negative ions on the P side of the junction and electrons being taken from the negatively charged ionized accepter atoms tends to neutralize the negative charge on the P side of the junction.

On the N side of the crystal, the negative terminal of the battery repels electrons towards the junction. These electrons tend to neutralize the positive charge on the donor atoms at the N side of the junction. At the same time the negative potential at the N side of the crystal attracts holes away from the charged positive ions on the N side of the junction. Both of these actions tend to neutralize the positive charge on the donor atoms at the junction.

The effect of the battery voltage is to reduce the potential barrier at the junction and allow more majority carriers to cross the junction. This means that we will have more electrons flowing from the N-type material across the junction to the P-type material and to the positive terminal of the battery and more holes traveling from the P-type material across the junction to the N-type material and towards the end of the crystal connected to the negative terminal of the battery. You know that we already had a certain number of intrinsic minority carriers crossing the junction, but now the majority carriers outnumber them, so there will be a steady current flow from the negative battery terminal, through the N-section, across the junction and through the P-section, to the positive battery terminal.

Placing a forward bias on a junction diode drives majority carriers back into the depletion layer and allows conduction across the junction. If the battery voltage is increased, more carriers will arrive at the junction and the current flow will increase. Eventually, if we continue to increase the battery voltage, we will reach a point where all the charges at the junction are neutralized. When this happens, the holes will fill the P-type region right up to the junction; electrons will fill the N-type region up to the junction; and the only limit to current flow through the diode will be the resistance of the material on the two sides of the junction.

It is important for you to remember that in a forward biased junction conduction through the crystal will be by the majority carriers. Any intrinsic conduction across the junction will be by minority carriers and this will subtract from the total current flow across the junction. Increasing the forward bias will increase the current flow across the junction until the point is reached where all the charges at the junction are neutralized, at which time the potential barrier will disappear, and current flow across the junction will be unhindered by any potential across the junction.

Reverse Bias.

If we reverse the battery connections we will have what is known as reverse bias. This condition is shown in Fig. 15.

With a reverse bias applied to a junction diode, the negative terminal of the battery will be connected to the P-type section, and will attract holes away from the junction, and increase the shortage of holes on the P side of the junction. At the same time the positive terminal of the battery is connected to the N-type section of the crystal and this

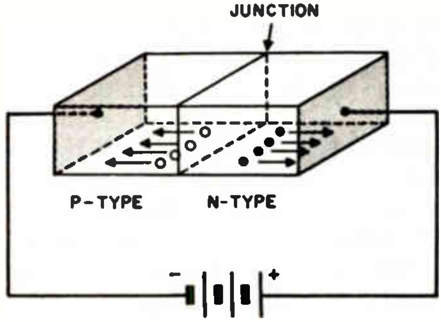


Fig. 15. Reverse-biased junction.

terminal will attract electrons away from the junction and increase the shortage of electrons on the N side of the junction. This movement of holes and electrons away from the junction will in effect result in an increased potential barrier at the junction. The increase in potential barrier occurs because there will be fewer holes on the P side of the junction to neutralize the negative ions and fewer electrons on the N side to neutralize the positive ions formed on this side of the junction. The increase in potential barrier will help prevent any further current flow across the junction due to majority carriers.

The current flow across the barrier, however, is not zero because we will still have minority carriers crossing the junction. Holes forming in the N side of the depletion layer will be attracted by the negative potential applied to the end of the P-type section of the crystal, and electrons breaking loose from their nuclei in the P side of the depletion layer will be attracted by the positive voltage applied to the end of the N-type section of the crystal.

We had this situation when there was no bias applied to the junction. Holes from the N side would cross over to the P section, and electrons from the P side would cross over to the N section. However, when there was no bias applied to the crystal, these minority carriers would neutralize ions near the junction and allow the majority carriers to cross the junction. However, since the minority carriers are now attracted away from the junction by the potential applied to the crystal, all of the minority carriers do not remain near the junction to neutralize charged atoms so they no longer allow the passage of an equal number of majority carriers in the opposite direction. This means that the flow of minority carriers across the junction is not fully offset by a flow of majority carriers in the opposite direction. Therefore, there will be a small current flow across the junction due to the minority carriers crossing the junction. This current flow is very small and nearly constant at all normal operating voltages in signal diodes and power rectifier diodes. However, as you will see later, there are certain types of diodes where this reverse current can increase quite rapidly even at low voltages.

It is important to realize that when a reverse bias is applied to a junction diode, the bias increases the potential difference across the junction and makes it more difficult for majority carriers to cross the junction. However, some minority carriers will still cross the junction with the result that there will be a small current flow across the junction due to the minority carriers.

COMPARISON OF JUNCTION DIODES AND VACUUM TUBES

Although the operation of junction diodes designed for use as rectifiers is quite different from the operation of vacuum tubes, they can perform identical tasks and therefore some comparison of the most important characteristics of both is in order.

When there is no voltage applied to a junction diode, the net current flow across the junction is zero. On the other hand, in a vacuum tube, even though there may be no voltage applied to the plate of the tube, some of the electrons will leave the cathode with sufficient velocity to travel across the space between the cathode and the plate and strike the plate. This will result in a small current flow from the cathode to the plate of the tube even though there may be no voltage applied to the plate of the tube.

We can consider applying a positive voltage to the plate of a vacuum tube and a negative voltage to the cathode as being similar to placing a forward bias on a junction diode. Under both circumstances there will be a current flow through the diode. In this respect the two are similar.

When the voltages applied to the diode vacuum tube are reversed so that there is a negative voltage applied to the plate and a positive voltage to the cathode, there will be no current flow at all through the tube. The negative potential on the plate of the tube will repel electrons from the plate. This reverse voltage situation is similar to a reverse bias across a junction diode. However, when we place a reverse bias across a junction diode, there will be some current flow across the junction due to the conduction by minority carriers. As long as the breakdown voltage of the junction diode is not exceeded, this current will be very small and almost constant. In a good diode, it is so small it can be ignored.

We can summarize the characteristics of diode vacuum tubes and junction diodes as follows: with forward bias both the tube and junction diodes will conduct. With reverse bias, the tube will not pass current; the junction diode will pass a small current. With no bias, the tube will pass a small current; the junction diode will not.

ZENER DIODES

In junction diodes designed for use as rectifiers we must be careful not to exceed the rated reverse voltage of the diode. In other words, if we place too high a reverse bias across the junction, the junction will break down, a very high current will flow across the junction for a short while, and the diode will be destroyed. However, in some diodes we make use of this reverse current due to minority carriers. In diodes of this type both the P section and the N section are doped quite heavily. The junction between the P section and the N section is considerably larger than the junction of the rectifier-type diode, so that when the diode begins to pass current in the reverse direction, it can pass it over a larger area and thus avoid destroying the diode. This type of diode is used as a voltage reference and is referred to as a voltage-reference diode or a Zener diode.

In the Zener diode, the current remains small with low reverse voltages. At a certain voltage, called the breakdown voltage, the current will increase rapidly with any further increase in voltage.

The breakdown voltage can be varied by varying the diode material and construction. Zener diodes can be made with a breakdown voltage as low as 1 volt, up to breakdown voltages of several hundred volts. The current that can pass through any Zener diode before the diode will be damaged will depend upon the junction area and the methods used to keep the diode cool.

In a circuit where a Zener diode is used, it will be used as a voltage reference or a voltage regulator. As the reverse voltage across the diode

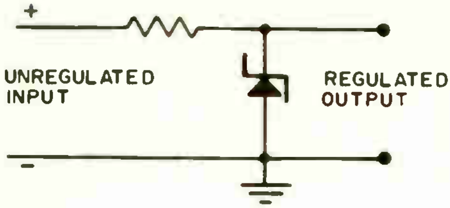


Fig. 16. Circuit using Zener diode as a voltage regulator.

increases, a very small reverse current will flow, and the value of this current will remain essentially constant until the breakdown voltage is reached. At that voltage, any further increase in voltage will result in a large increase in current across the junction. This large increase in current tends to produce a voltage drop in other components in the circuit so that the voltage across the diode will remain essentially constant. Thus a Zener diode can be used as a voltage regulator in a circuit such as shown in Fig. 16. If the input voltage tends to rise, the current through the diode will increase. This increase in current will increase the voltage drop across the resistor so that the output voltage will remain essentially constant. The fact that with even a very small increase in voltage the current through the diode will increase substantially means that the voltage across the diode will remain almost constant.

In some other applications the diode may be used as a reference voltage. This simply means that a Zener diode with a given breakdown voltage is used in a circuit like Fig. 16. The voltage applied to the diode will remain essentially constant. In some circuits, we may compare the voltage across a resistor or another part with the voltage across the Zener diode.

TUNNEL DIODES

A tunnel diode is a highly doped junction diode made either of germanium or gallium arsenite. Both the N region and the P region of the diode are very highly doped. As a result of the high doping, the depletion region around the junction is extremely narrow. Because of the narrow depletion region, holes and electrons can cross the junction by more or less tunneling from one atom to another. The exact action of the charges crossing the junction is somewhat difficult to visualize, but the characteristics of the tunnel diode are comparatively simple. With a reverse bias across the junction, current across the junction increases quite rapidly. If the reverse bias is dropped to zero the current will drop back to zero. If the forward bias, starting at zero and gradually increasing, is placed across the junction, current flow across the junction will increase at essentially the same rate as it increased with a negative bias. The current across the junction will increase quite rapidly as the forward bias is increased until a rather sharp peak is reached. If the forward bias is increased beyond this point, then the current begins to decrease.

As the forward bias is increased still further, the current across the junction decreases, forming a curve such as shown in Fig. 17. This decreasing current, with increasing voltage, results in a negative resistance characteristic. It might be difficult to visualize what a negative resistance is, but you will remember from Ohm's Law that re-

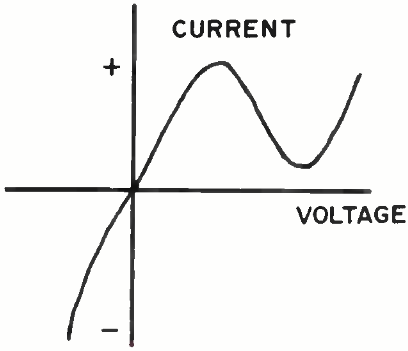


Fig. 17. Voltage-current relation in a tunnel diode.

sistance is equal to voltage divided by current. In a circuit where we have resistance, if the resistance is constant and the voltage increases, then the current must increase; and similarly if the voltage decreases, the current must decrease. Here in the tunnel diode we have a region where the opposite happens. If the voltage increases, the current decreases and if the voltage decreases, the current increases. Thus we have something in the circuit that is giving us the opposite effect of resistance; we call this negative resistance. You will remember that resistance in a circuit introduces losses. It is the resistance in a resonant circuit that prevents a resonant circuit from continuing to oscillate once it has been excited into oscillation. However, if we can put something with negative resistance in the circuit, for example a tunnel diode, since it has the opposite effect of resistance, then the circuit should continue to oscillate. Tunnel diodes can be used for this purpose.

At the present time, tunnel diodes have not appeared in the commercial entertainment-type equipment. However, it is probably just a matter of time until they are used; therefore you should at least have some basic knowledge of what the tunnel diode is.

P-I-N DIODES

The p-i-n diode, which is an abbreviation from positive-intrinsic-negative, is a new diode which is used in a somewhat different manner from the diodes you have studied previously. Rather than being used as a detector or rectifier, this diode is used primarily as a variable resistor. It is a special type of diode, and its resistance can be controlled by applying a dc bias to it. With a reverse bias across the diode, it has a very high resistance. With no bias, its resistance drops to about 7000 ohms, and with forward bias it drops to a comparatively low value.

The diode is particularly useful in circuits where the strength of a signal must be controlled. Its first commercial use has been in fm equipment and it is used in order to prevent extremely strong fm signals from causing overloading in the fm receiver. A dc bias is applied to the diode and the amplitude of the bias depends upon the strength of the signal. When a very strong signal is applied, the reverse bias applied to the diode increases so that the resistance of the diode increases. This reduces the strength of the signal fed to the mixer and i-f stages in the receiver and thus prevents overloading, particularly in the last stage of the receiver.

At this time p-i-n diodes are not widely used in commercial applications, but you should be aware of how the diode is used, because it is quite likely that it will be widely used in the future.

POINT-CONTACT DIODE

Another semiconductor diode is the diode detector used in many TV receivers. This detector is a point-contact diode. A cut-away view of a point-contact diode is shown in Fig. 18 along with the schematic symbol.

The point-contact diode is made of a small piece of either N-type or P-type germanium or silicon. N-type germanium or silicon is used more often than P-type. In manufacturing a diode of N-type material, a large contact is fastened to one side of the crystal. A thin wire, called a catswhisker, is attached to the other side of the crystal. When the catswhisker is attached to the N-type crystal, a small region of P-type material is formed around the contact as shown in Fig. 19. Thus we have a PN junction that performs in much the same way as the junctions we have already discussed.

The characteristics of the point-contact diode under forward and reverse bias are somewhat different from those of the junction diode. With forward bias the resistance of the diode is somewhat higher than that of a junction diode. With reverse bias the current flow through a point-contact diode is not as independent of the voltage applied to the crystal as it was in the junction diode. In spite of these disadvantages, the point-contact diode makes a better detector than the junction diode, particularly at high frequencies because the point-contact diode has a lower capacity than the junction diode.

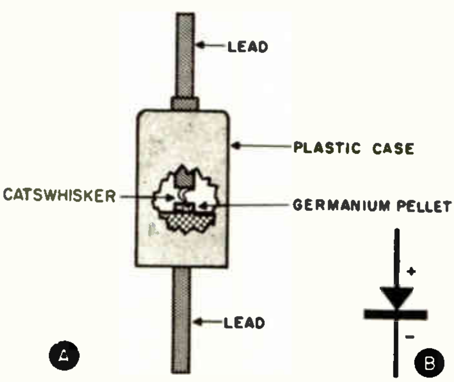


Fig. 18. (A) cut-away view of crystal rectifier. (B) schematic symbol of it.

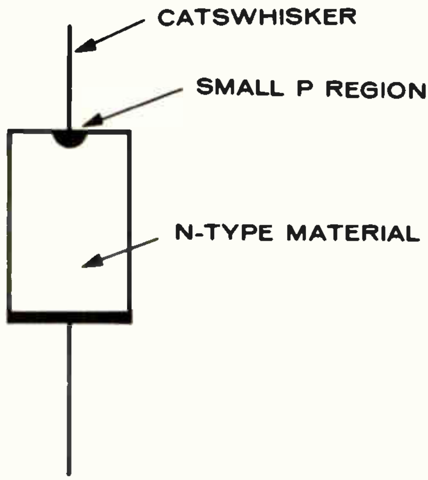


Fig. 19. Sketch of a point-contact diode showing where the P-type germanium is formed around the catswhisker.

SUMMARY

Again this is a very important section, so you should review it before going on to the next section. Make sure you understand what the depletion layer is and why it is formed. Also be sure you understand the movement of holes and electrons across a PN junction when no voltage is applied to the junction.

Current flow across the junction with both forward and reverse bias is important. With forward bias current flow across the junction is by majority carrier, and with reverse bias it is by minority carrier.

SELF-TEST QUESTIONS

(ab) What are the two principal uses of semiconductor diodes?
(ac) What is the depletion layer?
(ad) What do we mean by the potential barrier?
(ae) Does the crystal develop an overall charge as a result of diffusion across the junction?
(af) Do the minority carriers crossing the junction have any adverse effect on the diode?
(ag) What do we mean when we say the junction is forward biased?
(ah) What do we mean when we say a junction is reverse biased?
(ai) What is the difference between vacuum-tube diodes and semiconductor diodes insofar as current flow through the diode is concerned when no voltage is applied?
(aj) What is the difference between current flow in a semiconductor diode and a vacuum tube under reverse voltage conditions?
(ak) What is a Zener diode, and what is it used for?
(al) What is a tunnel diode?
(am) What is a p-i-n diode?

Semiconductor Triodes

Even though a junction diode will pass current in both directions, it passes current in one direction much better than it does in the other, and therefore it can be used as a detector. The tunnel diode can be used as an oscillator, and in some special circuits as an amplifier; however, its usefulness in these applications is limited. In most cases, the semiconductor diode is like the vacuum-tube diode; it is more or less useless insofar as amplifying a signal is concerned. In order to amplify a signal, a three element semiconductor is needed. Three element semiconductors that are capable of amplification are called transistors.

There are a number of different types of transistors in use today. The characteristics of the different types vary appreciably, but if you understand the operation of one type, you can understand how the others work without too much difficulty. We started our explanation of semiconductor devices with a junction diode, so we will start our study of triode semiconductors with a study of the junction transistor. You'll find that most of the transistors you will study operate in a manner similar to the basic junction transistor. The most notable exception to this is the field-effect transistor which you will study later in this lesson.

JUNCTION TRANSISTORS

Both germanium and silicon are used in the manufacture of junction transistors. A triode-junction transistor is made up of single semiconductor crystals with three different regions. The center region is made up of one type of germanium or silicon, and the two end regions are made up of the other type of germanium or silicon. In other words, in one type of junction transistor the center has had accepter-type region impurities added and the two end regions have had donor-type impurities added. In the other type of junction transistor, the center region has had donor-type of impurities added and the two end regions have had accepter-type impurities added.

The center region of the transistor is called the base. This is usually a comparatively thin region. One of the end sections is called the emitter and the other end section is called the collector.

If the center section of the crystal has been treated with donor-type impurities, then the center section becomes N-type germanium in the case of a germanium transistor or N-type silicon in the case of the silicon transistor. In this case, the two end sections will be treated with accepter-type impurities and they will both become P-type germanium or P-type silicon. We call this type of transistor a PNP transistor. We can have both germanium PNP and silicon PNP transistors. An example of this type of junction transistor is shown in Fig. 20A along with the schematic symbol used to identify it.

The other type of junction transistor is an NPN type. This type of transistor is produced by treating the center section with an accepter-

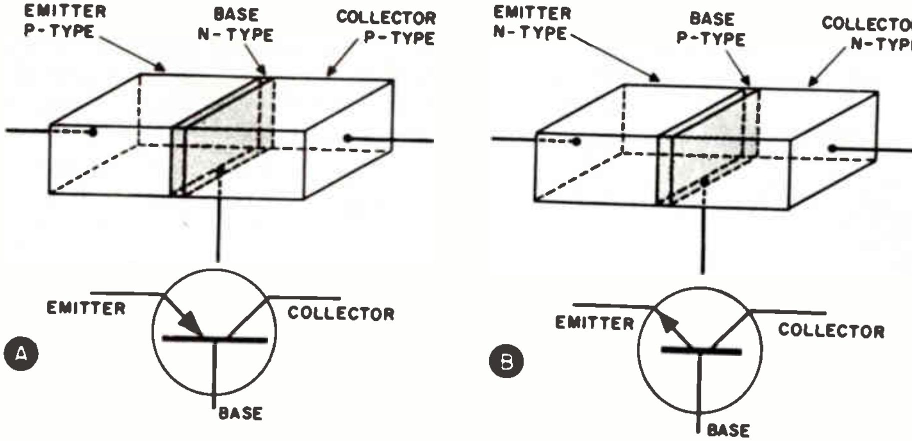


Fig. 20. (A) shows a PNP junction transistor and its schematic symbol. (B) shows an NPN junction transistor and its schematic symbol.

type impurity to produce P-type germanium or silicon, and the two end sections with a donor-type of impurity to produce N-type germanium or silicon. This is the type of junction transistor shown in Fig. 20B along with the schematic symbol for it. As in the case of the PNP transistor, we can have either an NPN germanium transistor or an NPN silicon transistor.

Notice that the schematic symbols for the PNP transistor and the NPN transistor are different. In the PNP transistor the arrow used to represent the emitter points down toward the base, whereas in the NPN transistor the arrow on the emitter points up away from the base. Thus on a schematic diagram of a piece of equipment using junction transistors, you can tell from the direction in which the arrow is pointing whether the transistor is a PNP or an NPN transistor.

There are several different ways of manufacturing junction transistors, and often you hear these transistors referred to by the manufacturing method used. For example a junction transistor might be called a grown junction, a fused junction, an alloy junction or a diffused junction. All these names simply describe the manufacturing process used to make the transistor. They are all junction transistors and all operate on the same basic principle. This does not mean to imply that the characteristics of all junction transistors are the same; they are not. There are wide differences in characteristics just as there are in various triode vacuum tubes.

To understand how junction transistors work, you must understand junction diodes. Because it is so important that you understand the formation of the depletion layers at the junctions we will review the explanation of what happens at the junctions in explaining the junction transistors. The big difference between the junction transistor and the junction diode is that in the transistor there are two junctions close together in the same crystal. One of these junctions is biased in the forward direction and the other in the reverse direction and the presence of one junction affects the operation of the other.

PNP TRANSISTORS

In transistor operation, the emitter-base junction is always biased in the forward direction and the collector-base junction is biased in the reverse direction. Each of the two junctions by itself behaves just like the PN junction already described.

Let us consider what happens in the PNP transistor before any voltages are applied to the transistor.

At the junctions, holes from the P-type emitter section and the P-type collector diffuse across the junctions into the base. At the same time, electrons from the base diffuse across the junctions into both the emitter and the collector. The holes diffusing into the base place a positive charge on the atoms near the junctions. Similarly the electrons diffusing from the base into the emitter on one side of the base and the collector on the other side of the base place a negative charge on the atoms on the emitter and collector sides of the junctions. These charged atoms, which are called ions, will repel electrons and holes from the region of the junctions. The positively charged ions in the base will repel holes in the P sections away from the junctions. Similarly, the negatively charged ions in the emitter and collector will repel electrons away from the junctions in the base. Thus we have two depletion layers formed, one at the emitter-base junction and the other at the base-collector junction.

You will remember that when we discussed the junction diode, we mentioned that hole-electron pairs will be formed in the depletion region. The minority carriers formed in each section can cross over the junction. For example, the electrons released in both the emitter and the collector regions will cross the junctions into the base. These electrons will neutralize a few ions in the base region. When these ions are neutralized they will allow majority carriers from the emitter and collector to cross the junctions. In other words, there will be holes from the emitter and holes from the collector crossing the junctions into the base. Similarly, holes, which are the minority carriers in the base region (and are formed in the depletion layer) will cross the junctions into the emitter and collector. When these holes cross the junctions they will neutralize some of the nega-

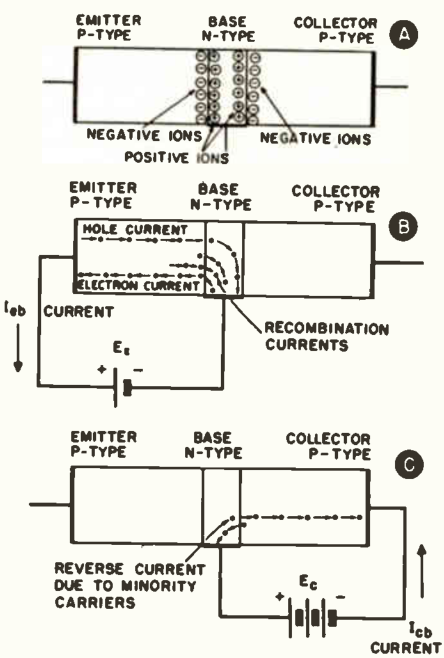


Fig. 21. (A) the formation of ions at the junctions of a PNP transistor. (B) the current flow in the emitter-base circuit and (C) in the collector-base.

tively charged ions in the emitter or collector region and allow some electrons to flow from the base into either the emitter or the collector. Thus, because of the intrinsic conduction due to hole-electron pairs being formed in the depletion region there will be some flow of carriers across the junction. However, the flow of majority carriers across the junction will be exactly equal to the flow of minority carriers across the junction so that the net current flow across each junction will be zero.

The potential barriers formed at the junction regions are shown in Fig. 21A. Notice that the charges formed at the junction are similar to those formed at a junction diode; we simply have two junctions to consider in a transistor.

Now when we place a forward bias between the emitter and the base we have an arrangement like that shown in Fig. 21B. Here the positive voltage applied to the end of the P-type emitter repels holes towards the junction. These holes tend to neutralize the negative charge on the ions on the emitter side of the junction. The holes are formed at the end of the P-type section by electrons being taken out of this section by the positive potential applied to it. At the same time the positive potential applied to the emitter attracts the electrons that have given the ions on the P side of the junction their negative charge. This also weakens the negative charge on the emitter side of the junction.

At the base, which is connected to the negative side of the battery, the holes will be attracted toward the negative terminal of the battery, and electrons will be pushed towards the depletion layer. The pulling of holes away from the depletion area and pushing electrons into the depletion area tends to neutralize the charge on the base side of the junction.

The net effect of biasing in a forward direction is to neutralize the charges on each side of the junction and allow current to flow across the junction. Current flow is by majority carriers: electrons from the N-type base region and holes from the P-type emitter region.

Thus in the emitter-base circuit we have electrons flowing from the negative terminal of the battery to the base, through the base, across the junction, and through the emitter to the positive terminal of the battery. At the same time we have holes being produced because electrons are being pulled out of the P-type emitter by the positive potential applied to it. The holes will move through the emitter, across the junction into the base and to the point where the base is connected to the negative terminal of the battery. At this point they will pick up electrons and disappear.

Not all the electrons going from the base to the emitter will reach the positive terminal of the battery. Some of these electrons will recombine with holes in the emitter. Similarly, some of the holes traveling from the emitter into the base will pick up an electron in the base. This current flow across the junction is called a recombination current, and the transistor is designed to keep this current as low as possible. In other words we want the holes and electrons crossing the junction to reach the terminals connected to the battery.

Now let us consider the other junction, the base-collector junction. This junction is reverse biased as shown in Fig. 21C. Here again we have a depletion layer at the junction. Also we have minority carriers being formed in the depletion layer. However, holes that are formed in the base will cross the junction and then instead of neutralizing a negatively charged atom near the junction in the collector, these holes will be attracted by the negative potential applied to the collector. Similarly, electrons formed in the depletion layer of the P-type collector will cross the junction and be attracted by the positive potential applied to the base. Thus we have a current flow due to the minority carriers. Electrons in the depletion layer of the collector section will cross the junction and flow through the base to the positive terminal of the battery. Meanwhile electrons from the negative terminal of the battery will fill holes that are moving from the base, across the junction, and through the collector to the negative terminal.

Thus you can see that while we have a current by majority carriers due to the forward bias applied between the emitter and the base, we also have a small current flowing through the base-collector circuit by minority carriers due to the reverse bias applied between the base and the collector. Now let us see how the two junctions affect each other.

A transistor with both biased junctions is shown in Fig. 22. Here we have a number of different currents flowing. In the emitter-base circuit we have current flowing due to the forward bias applied between these two. Electrons will flow from the negative terminal of the battery into the base, across the junction, and through the emitter to the positive terminal of the battery. We will also have some holes formed in the P-type emitter section due to electrons being pulled out of this section by the positive terminal of the battery. Some of these holes will cross the junction into the N-type base where they will pick up an electron and disappear. This current is called the recombination current.

Many of these holes will cross the base and flow through the collector, because the negative terminal of the battery connected between the base and collector will attract them. This movement of holes accounts for most of the current flow in the emitter and collector circuits. Remember that holes are being continually formed in the P-type emitter because electrons are

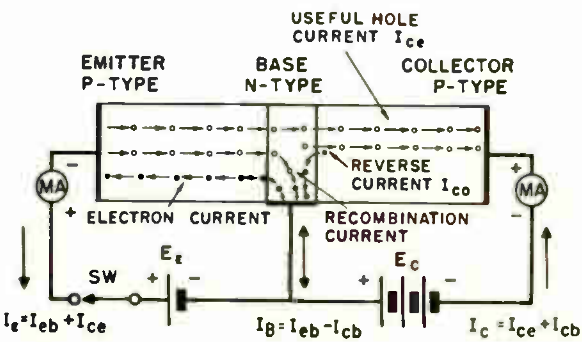


Fig. 22. Current flow and carrier movement in a PNP junction transistor.

being pulled out of the emitter by the positive potential applied to it. These holes will continually move through the emitter, and into the base. Here some of them will combine with electrons and disappear, but the majority of them will flow through the collector to the negative terminal of the collector where they will be filled by electrons and disappear.

Another current that will flow is reverse current I_{CO} that flows in the base-collector circuit. This is due to the formation of minority carriers in the depletion layer.

Thus we have four currents flowing in the PNP junction transistor. The largest of these currents is due to the movement of holes from the emitter through the base into the collector to the negative terminal of the battery connected to the collector. We have in addition to this current three small currents flowing. We have the current due to the electron movement from the negative terminal of the emitter-base battery into the base, across the junction and through the emitter to the positive terminal of this battery. We have the recombination current due to holes combining with electrons in the base, and we have the reverse current due to hole-electron pairs being formed in the depletion layer of the base-collector junction. The directions of the different movements of holes and electrons are marked in Fig. 22.

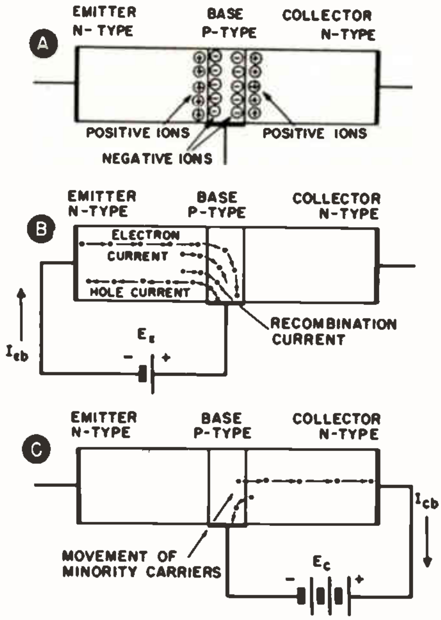


Fig. 23. (A) the formation of ions at the junctions of an NPN transistor. (B) the current flow in the emitter-base circuit and (C) in the collector-base circuit.

NPN TRANSISTORS

Although the operation of the NPN transistor is somewhat different from that of the PNP type, if you understand how the PNP transistor works, you should have no difficulty understanding the NPN. Again, we can start our study of this type of transistor by considering what happens at the junctions, remembering that the action at the junction is similar to the action we studied at the simple PN diode junction.

Let us first consider the action of the holes and electrons before any voltages are applied to the transistor. The charges that will be built up are shown in Fig. 23A. Remember that holes from the base will diffuse across both junctions into the emitter and the collector. Similarly electrons from the emitter and electrons from the collector will diffuse across the junctions into the base. The holes and electrons diffusing across the junctions will charge atoms near the junction. Holes crossing the junctions into the emitter and the collector will ionize the atoms on the emitter and collector sides of the junctions so that they will have positive charges. Similarly, electrons diffusing across the junctions into the base will ionize atoms in the base near the junctions so that they will have negative charges. Thus we will have potential barriers at the junctions. This is the same kind of potential barrier that we found existed across the PN junction in a diode.

The positively charged ions on the emitter and collector sides of the junctions will force holes in the base away from the junction. Similarly the negatively charged atoms on the base side of the junctions will force electrons in the emitter and collector away from the junction so that at the junctions we have a depletion layer.

Now let us consider what happens when we apply a forward bias between the emitter and the base by connecting a battery between the two as shown in Fig. 23B. Notice that the negative terminal of the battery is connected to the end of the emitter, and the positive terminal is connected to the base.

Now, several things happen. The negative potential applied to the emitter will force electrons toward the junction. At the same time the negative potential will attract holes away from the junction. Both of these actions tend to neutralize the positively charged ions on the emitter side of the junction. At the same time the positive terminal of the battery that is connected to the base will attract electrons away from the negatively charged atoms on the base side of the junction. In addition, the positive potential will repel holes towards the junction so that these two actions tend to neutralize the charge on the base side of the junction.

Once the potential barrier at the junction is weakened, electrons can flow from the negative side of the battery into the emitter, through the emitter, and across the junction into the base and from the base to the positive side of the battery. At the same time the positive terminal of the battery can extract electrons from the base, forming holes. Holes are then repelled toward the junction, across the junction, and through the emitter toward the end of the emitter that is connected to the negative terminal of the battery. Here the holes will pick up electrons and disappear. Thus we have a current flow through the emitter-base circuit as shown in Fig. 23B.

Now let us consider what happens when we apply a reverse bias between the base and the collector. Here the negative potential applied to the base will pull holes away from the junction, and the positive potential applied to the collector will pull electrons away from the junction. Thus the negative charge on the base side of the junction will be increased, and the positive charge on the collector side of the junction will be increased so that the potential barrier at the junction will be increased. This will prevent any current flow through the base-collector circuit due to the majority carriers.

At the same time electrons, which are minority carriers, will break loose from their nuclei in the depletion layer on the base side of the junction and will be attracted by the positive potential applied to the collector. They will cross the junction and flow through the collector to the terminal connected to the positive side of the battery, as shown in Fig. 23C. At the same time holes formed on the collector side of the junction in the layer will be attracted by the negative terminal of the battery, and hence will cross the junction and flow over into the base and toward the negative terminal of the battery. Here they will pick up an electron and disappear.

Thus we will have a current flow in the base-collector circuit due to the minority carriers. This is the same situation that we had in the reverse biased base-collector circuit of the PNP transistor.

Now let us see what happens when bias voltages are applied across both junctions of the complete NPN junction transistor as shown in Fig. 24. Considering first the emitter-base circuit, we have electrons flowing from the negative terminal of the battery to the N-type emitter. Here the electrons flow through the emitter, across the junction, and into the base. Some of these electrons reaching the base will recombine with holes in the base. This is called the recombination current. However, the majority of the electrons reaching the base will be attracted by the positive potential applied to the collector and hence will flow through the base across the base-collector junction and through the N-type collector to the positive terminal of the battery in the base-collector circuit.

At the same time the positive terminal of the battery in the emitter-base circuit is connected to the base, and this potential will pull electrons out of the P-type base, producing holes. These holes will then cross the junction into the emitter, and they will be attracted by the negative potential applied to the emitter and hence will flow through it to the end connected to the negative terminal of the battery. Here they will pick up an electron and disappear.

At the same time, in the base-collector circuit we will have a reverse current flowing due to the minority carriers. Holes appearing in the collector side of the depletion layer will cross the junction into the base and flow to the base terminal connected to the negative terminal of the battery, biasing the base-collector junction. Here each hole will pick up an electron and disappear. Electrons in the depletion layer on the base side of the junction will be attracted by the positive potential applied to the end of the collector. Hence they will cross the junction and flow toward the positive end of

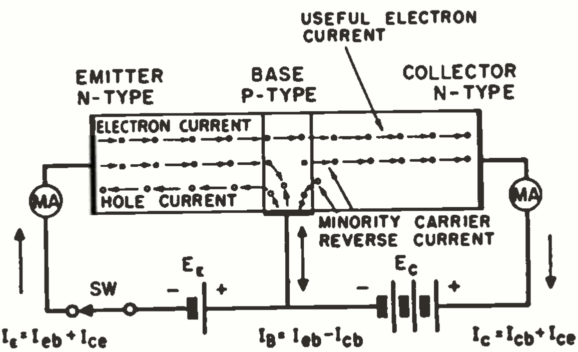


Fig. 24. Current flow and carrier movement in an NPN junction transistor.

the collector and from there to the positive terminal of the battery connected between the base and the collector.

Of these different currents flowing, the important and useful current flow is the flow of electrons from the emitter through the base to the collector. Since this is the useful current, we are interested in making this as large as possible in comparison to the other currents flowing across the emitter-base junction. Thus, the recombination current, which is due to electrons from the emitter crossing into the base and recombining with the holes, serves no useful purpose and should be kept as low as possible. This is accomplished by adding more donor atoms to the emitter than accepter atoms to the base. Thus there will be many more free electrons in the emitter than there will be holes in the base and the recombination current will be kept quite small.

Also, since there are a limited number of holes in the base compared to the number of electrons in the emitter, the number of holes crossing from the base to the emitter is also kept low in comparison to the number of electrons crossing from the emitter into the base. In a good transistor, over 95% of the electrons that cross the emitter-base junction flow to the collector.

Notice the differences and the similarities between the PNP and the NPN transistors. In both cases the emitter-base junction is forward biased and the base-collector junction is reverse biased. However, the battery connections must be reversed to provide the biases. In other words, with the PNP transistor the battery used to bias the emitter-base junction is connected with the positive terminal to the emitter and the negative terminal to the base. With the NPN transistor, the negative terminal of the battery is connected to the emitter and the positive terminal to the base. However, both are forward biased because in each case the positive terminal of the battery is connected to the P-type germanium and the negative terminal to the N-type germanium.

The base-collector junction of both transistors is reverse biased. In the PNP transistor, the positive terminal of this battery is connected to the base and the negative terminal to the collector; whereas in the NPN transistor, the negative terminal is connected to the base, and the positive terminal to the collector. Again, however, in both cases the positive terminal is connected to the N-type germanium and the negative terminal to the P-type germanium.

Also notice that in the PNP transistor the useful current flow is by means of holes, whereas in the NPN transistor the useful current flow is by means of electrons.

SELF-TEST QUESTIONS

(an) What is the base region of a transistor?

(ao) What two materials are widely used in the manufacture of transistors?

(ap) What two types of junction transistors are widely used?

(aq) What type of bias is used across the emitter-base junction in a transistor?

(ar) What type of bias is used across the base-collector junction of a transistor?

(as) Is the base region of a transistor usually a thick region or is it thin?

(at) Draw a diagram of a PNP transistor and show how the batteries are connected to place the correct bias across the two junctions.

(au) Draw a diagram of an NPN transistor and show how the batteries are connected to provide the correct bias across both junctions.

(av) What are the useful current carriers in a PNP transistor?

Semiconductor Types

There are two basic types of transistors that you will run into continuously. You are already familiar with these two types; they are the NPN transistor and the PNP transistor. However, these transistors are made in a number of different ways and the manufacturing processes result in transistors with different characteristics. In this section we are going to briefly discuss some of the important types and characteristics. We don't expect you to remember all these details; the important thing for you to remember is that they are basically either NPN or PNP transistors and operate in the same way as those we have discussed previously.

Also in this section of the lesson we'll discuss two other important semiconductor devices, the field-effect and unijunction transistors.

GROWN-JUNCTION TRANSISTORS

The first commercially available junction transistors were of the grown-junction type. This type of transistor is made from a rectangular bar cut from a germanium crystal that has been grown. Suitable impurities are added so that NPN regions such as those shown in Fig. 25 are formed. The base of the transistor is usually located midway between the two ends. Suitable contacts are then welded to the emitter, base and collector regions.

Of course, the actual bar of semiconductor material used is quite small. The emitter and the collector are considerably larger than the base; the base is kept as thin as possible and may have a thickness of less than .001".

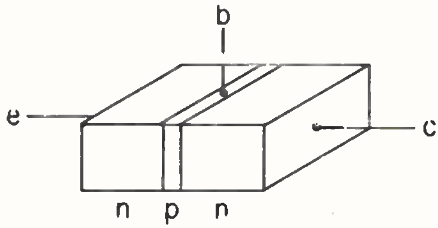


Fig. 25. A grown-junction transistor.

As mentioned the early germanium transistors were of the grown-junction type. The disadvantage of this type of transistor is that it is not particularly suitable for operations at high frequencies. In addition, it is quite temperature sensitive and can become quite unstable at higher temperatures.

ALLOY-JUNCTION TRANSISTORS

The alloy-junction transistor is made from a rectangular piece of

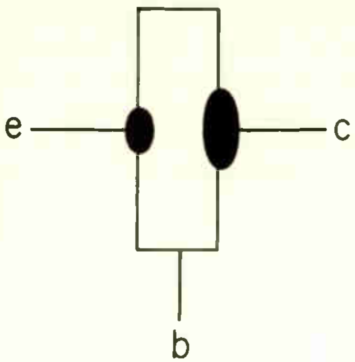


Fig. 26. An alloy-junction transistor.

semiconductor material to which suitable donor materials have been added. This results in an N-type piece of germanium or silicon. Small dots of indium are fused into the opposite sides of the wafer as shown in Fig. 26. The result is that P-type semiconductor material will be formed with the dots fused into the wafer so that we will have a PNP transistor.

An NPN-type alloy-junction transistor may be made by fusing a lead antimony alloy into each of the two opposite sides of a P-type semiconductor wafer. In this type of transistor it is possible to get a more uniform penetration of the lead antimony alloy into the semiconductor material, and this in turn leads to better junction spacing. This will cut down on the width of the space between the emitter and collector and give improved high-frequency performance. In addition, since the mobility of the electrons is more than twice that of holes, the NPN transistor will be better at high frequencies.

The general advantage of the alloy-type junction over the grown-type junction transistors is that they are usable at a somewhat higher frequency. In addition, they have a higher current gain, and the current gain remains stable as the temperature increases.

Surface-Barrier Transistor.

The surface-barrier transistor is similar to the alloy-type transistor except that depressions are etched into the N-type wafer. This permits smaller emitter and collector contacts and results in lower capacities between sections of the transistor which in turn results in better high-frequency performance.

In Fig. 27 we have shown a simplified sketch of a surface-barrier transistor. The sketch in Fig. 27B shows the carrier movement from the emitter across the base to the collector. Notice that in the sketch the emitter is shown smaller than the collector, we have shown it this way because this is the way the semiconductor is actually manufactured.

Various manufacturing techniques are used in the manufacture of the surface-barrier transistor. Both silicon and germanium types are made. In the manufacturing process different materials are evaporated or plated on to the etched depressions depending on the type of tran-

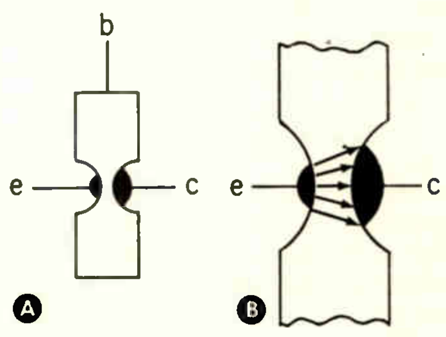


Fig. 27. Sketch of a surface barrier transistor is shown at A. Hole movement across the base is shown at B.

sistor being manufactured. However, regardless of the manufacturing technique used, which is of no interest to the technician, the surface-barrier transistors all have the characteristic of giving good performance at high frequencies.

DIFFUSION TRANSISTORS

To understand diffusion you have to understand a little about the molecular structure of materials. If you look at the wall of a glass jar, to the eye it appears solid with no space between the various molecules making up the jar. However, if you were to fill the jar with hydrogen and store it for any length of time, you would find that in a short while, the jar was no longer filled with hydrogen only, but contained a mixture of hydrogen and air. The reason is that the small hydrogen atoms are able to diffuse or pass right through the spaces between the molecules in the glass. At the same time, molecules of air will diffuse through the glass and pass on into the inside of the bottle. The hydrogen molecule is smaller than the air molecule; therefore the hydrogen will diffuse out of the jar faster than the air will diffuse in.

Diffusion can be used to add impurities to either silicon or germanium, and produce either N-type or P-type semiconductor material. The process can be controlled to provide either very uniform base, emitter, and collector regions, or it can be controlled to provide non-uniform base, emitter, and collector regions.

The Drift Type.

One of the most important uses of the diffusion technique is in the manufacture of transistors with a non-uniform base region. If the emitter and collector junctions are made by the alloy technique, but the base region is made by the diffusion technique and the impurities in the base region varied, we have what is known as a drift transistor. In a typical PNP-drift transistor, accepter impurities are added in the emitter and collector region. These impurities are controlled so that their concentration is uniform throughout the emitter and collector region. At the same time donor impurities are added to the base region. Their concentration is controlled so that it is highest in the region of the emitter-base junction and then drops off quickly and finally reaches a constant value which it maintains over to the base-collector junction, as shown in Fig. 28. This type of transistor is called a drift transistor, and its most important characteristic is its excellent performance

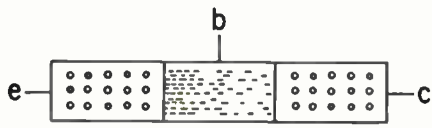


Fig. 28. Diagram showing how a large number of donor impurities increases the electron concentration in the base.

at high frequencies. However, notice that it is still a PNP transistor and the basic theory of its operation is similar to that of any other PNP transistor. The improved performance is obtained by varying the concentration of donor impurities in the base region.

The Mesa Type.

It is also possible to manufacture a transistor using the diffusion technique entirely. An example of this type is the mesa transistor.

In this type of transistor a semiconductor wafer is etched down in steps so that the base and emitter regions appear as plateaus above the collector region as shown in Fig. 29. The advantages of the mesa transistor are good high-frequency performance and very good consistency. By this we mean that it is possible to control the manufacturing techniques quite closely so that the characteristics of mesa transistors of the same type number will be quite similar. This is not necessarily true of other transistors; often their characteristics vary over a wide range.

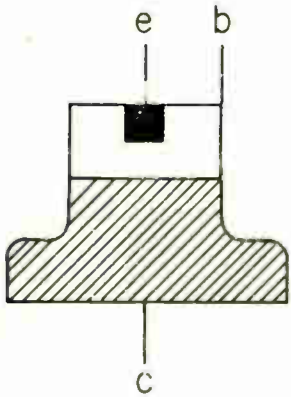


Fig. 29. A mesa transistor.

The Planar Type.

Another type of transistor manufactured by the diffusion technique is the planar type of diffused transistor. This type of transistor is shown in Fig. 30. Notice that each of the junctions is brought back to a common plane, whereas in the mesa type the various junctions are built up in plateaus. The importance of the planar-type transistor is that the junctions can be formed beneath a protective layer. As a result, many of the problems associated with other types of transistors having

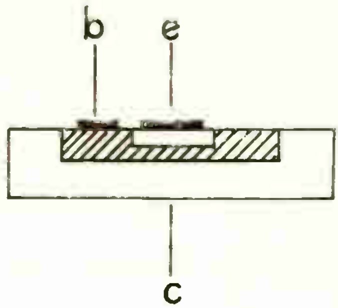


Fig. 30. A diffused planar-type transistor.

junctions exposed at the surface are avoided in this type of construction. Important characteristics of the planar transistor are generally very low reverse current and improved dc gain at low-current levels.

EPITAXIAL TRANSISTORS

One of the disadvantages of the diffusion-type transistor is the relatively high resistance of the collector region. This results in slow switching time; it limits the usefulness of the transistor in high-frequency applications. Reducing the resistance of the collector region reduces the collector breakdown voltage and this in turn again reduces the usefulness of the transistor. These problems can be overcome by the epitaxial technique. In this technique a thin high-resistance layer is produced in the collector region and the remainder of the collector region is controlled to keep its resistance low. This results in a transistor that looks something like the one shown in Fig. 31. The primary advantage of this transistor is that it provides good performance at very high frequencies. This technique can be combined with other techniques to produce transistors having varying characteristics. The epitaxial transistor can be referred to as a double-diffused epitaxial transistor. The thin high-resistance collector region is formed by the epitaxial technique and the base and the emitter are formed by the diffusion process - hence the term double diffusion.

All the transistors that we have discussed so far in this section of the lesson are either NPN or PNP transistors. The manufacturing techniques used to manufacture these transistors result in transistors of different characteristics, but the basic theory of operation of these transistors is the same. Now, we'll look at another semiconductor device which operates on a somewhat different principle.

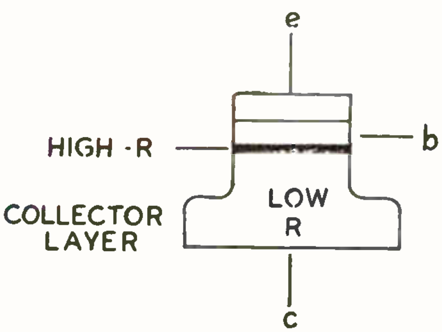


Fig. 31. A double-diffused epitaxial transistor.

THE JUNCTION FIELD-EFFECT TRANSISTOR

An interesting transistor that resembles a vacuum tube very closely in its characteristics and to some extent its operation is a field-effect transistor. One type of field-effect transistor can be made by taking a piece of N-type material as shown in Fig. 32. If the negative terminal of a battery is connected to one end of the material and the positive side

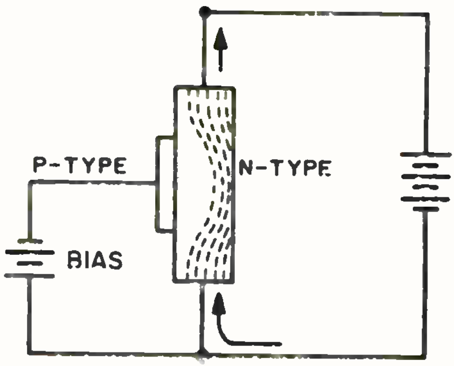


Fig. 32. Drawing showing the basic operation of a field-effect transistor.

of the terminal to the other end, electrons will flow through the material as shown. If we attach a piece of P-type material to one side so that the PN junction is formed and then place a negative voltage on the P-type material as shown in Fig. 32, there will be no current flow across the junction, because the battery biases the junction in such a way that electrons cannot flow from the N-type material to the P-type material nor can holes flow from the P-type material to the N-type.

However, the negative voltage applied to the P-type material sets up a field in the N-type material. This field opposes the electrons flowing through the N-type material and forces them to move over to one side so that the electron movement follows the path shown in Fig. 32. The negative voltage applied to the P-type material has the effect of increasing the resistance of the N-type material in the area in which the field is affected. It forms a depletion layer around the junction so there will be no free electrons in the N-type material near the junction. If the negative bias voltage is made high enough, it is able to prevent

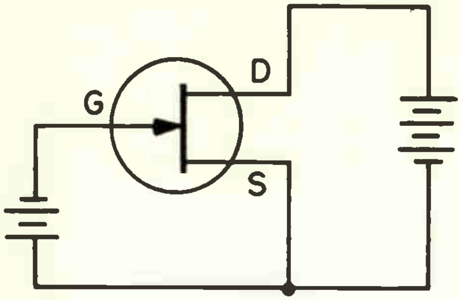


Fig. 33. Schematic representation of the circuit shown in Fig. 32.

the flow of electrons through the N-type material entirely so that the current flow will be cut off. We call this voltage where the bias voltage is high enough to stop the flow of current through the N-type material the "pinch-off" voltage. The N-type material is referred to as a channel, and the P-type material as a gate. This type of transistor is called a "junction field-effect transistor."

The schematic representation of the circuit shown in Fig. 32 is shown in Fig. 33. Notice that the end of the N-type channel at which the electrons from the battery enter is called the "source". The other end, the end from which the electrons leave and flow to the positive terminal of the battery, is called the "drain". The P-type material is called the gate, as we mentioned previously. The transistor is called a field-effect transistor because it is the field produced by the bias voltage applied to the gate that controls the flow of current through the channel. This particular type of transistor is called a junction transistor because a junction is formed between the P and N-type materials. It is called an N-channel transistor because the material in the channel through which current flows has been treated in such a way as to produce an N-type semiconductor material. Thus the complete name for this type of transistor is an N-channel, junction-gate, field-effect transistor. We usually abbreviate field-effect transistor FET, so you will see that this type of transistor is abbreviated JFET to indicate it is a junction-gate type.

An amplifier using a field-effect transistor of this type is shown in Fig. 34. In this circuit we have eliminated the bias battery by means of a resistor connected between the negative terminal of the battery and the source. This resistor might be compared to the cathode-bias re-

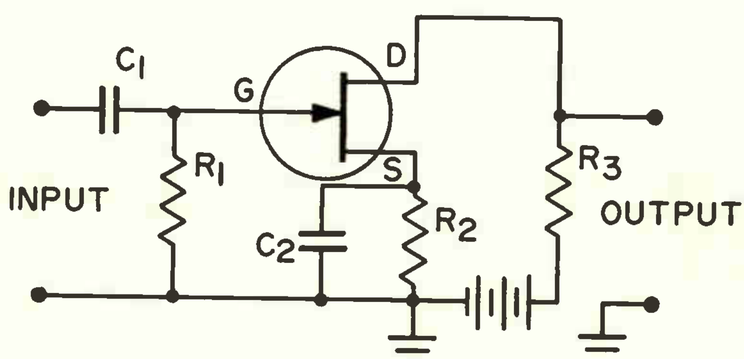


Fig. 34. An amplifier using an N-channel junction gate FET.

sistor in a triode vacuum tube amplifier stage. In the amplifier circuit, electrons flow from the negative terminal of the battery through the resistor R_2 to the source. In so doing they set up a voltage drop across R_2 having a polarity such that the source is positive with respect to ground. Since the gate connects back to ground through R_1, the gate will be at ground potential and this will make the source positive with respect to the gate, or in other words, the gate negative with respect to the source. Therefore none of the electrons in the N channel will flow to the gate, because the gate is negative.

Electrons will flow through the N channel to the drain and then through the load resistor R_3 back to the positive terminal of the battery. As the input voltage applied across the input terminals causes the voltage between the gate and the source to vary, the current flow from the source to the drain will vary because the controlling action of the gate on the current through the channel depends upon the voltage between the gate and the source. Thus we have a varying current, which will vary as the input signal varies, flowing from the source to the drain of the transistor and through the load resistor R_3. This varying current flowing through R_3 will produce an amplified signal voltage across R_3.

It is interesting to note the similarity between the circuit shown in Fig. 34 and a triode amplifier. When the input signal swings the gate in a positive direction, current flowing through the transistor will increase; this will cause the voltage drop across R_3 to increase and therefore the voltage between the drain and ground will decrease. Thus a positive-going signal applied to the gate will cause a negative-going signal at the drain. In other words, this transistor inverts the signal phase just as the triode vacuum tube amplifier stage does.

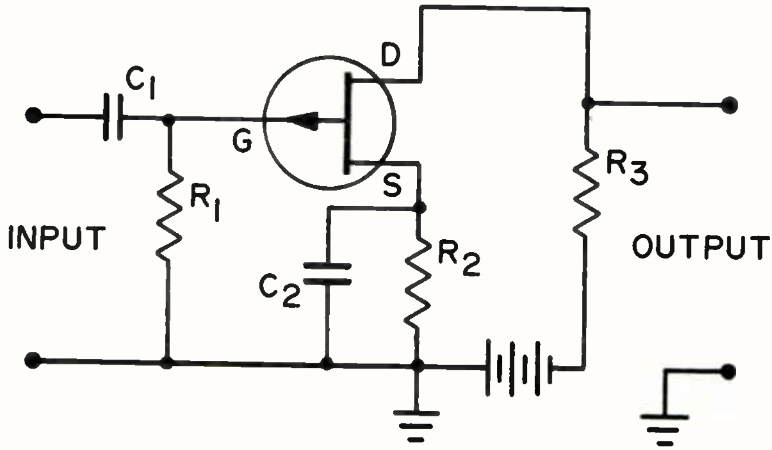


Fig. 35. An amplifier using a P-channel junction gate FET.

P-Channel JFET.

It is possible to make a P-channel junction-gate field-effect transistor by using a P-type material between the source and drain. The gate is then made of an N-type material. The bias polarity is reversed so that once again the PN junction is biased and no current flows across the junction.

A schematic diagram of an amplifier using a P-channel junction-gate effect is shown in Fig. 35. Notice the schematic symbol for the P-channel unit; we have turned the direction of the arrow around just as we did to distinguish between NPN and PNP transistors. Also notice that in this circuit the battery polarity is reversed. This is because the carriers in the channel in the P-channel unit will be holes. The positive terminal of the battery which connects to the source through R_2 repels the holes and they travel through the channel to the drain where they are attracted by the negative potential connected to the drain. Meanwhile, holes arriving at the drain terminal are filled by electrons which flow from the negative terminal of the battery through R_3 to the drain. At the same time, the positive terminal of the battery attracts electrons from the source creating new holes. These electrons flow from the source through R_2 to the positive terminal of the battery.

The operation of the P-channel, junction-gate effect is the same as with the N-channel unit, except that in one case the majority carriers are electrons, and in the other case they are holes.

In discussing the action of the junction-gate field-effect transistor, we often refer to the reverse bias across the junction creating a depletion layer in the conducting channel. In the case of an N-channel unit, the negative voltage on the P-type gate will repel electrons at the junction so that the electrons have been depleted from that area around the junction. The higher the negative voltage the further the electrons are depleted in the area around the junction, and as we pointed out previously if the voltage is made high enough, all of the electrons will be depleted so that there will be no current flow through the channel. The transistor is referred to as a depletion-type transistor because the bias depletes the number of majority carriers from the channel around the junction region. Remember what we mean by a depletion type of FET; you'll see later there is another type.

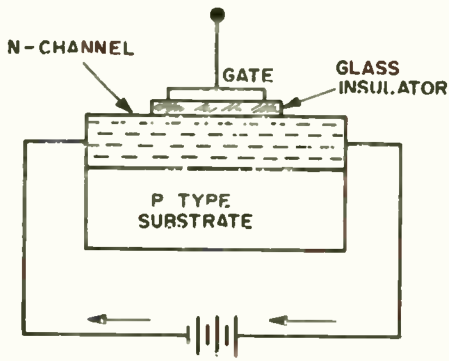


Fig. 36. Current flow through an insulated-gate, N-channel field-effect transistor with no bias applied.

INSULATED-GATE FIELD-EFFECT TRANSISTORS

The transistors we have been discussing so far are called junction-gate field-effect transistors. There is another type of field-effect transistor that is called an insulated-gate field-effect transistor. We usually abbreviate this IGFET.

In the insulated-gate field-effect transistor, the gate is completely insulated from the channel by a thin insulating material. For example, a very thin piece of glass might be placed between the conducting channel and the gate. Thus there is no actual junction formed between the semiconductor materials in the channel and the gate. In an N-channel, insulated-gate field-effect transistor, construction such as shown in Fig. 36 is often used. Here we have an N channel between the source and drain. The substrate on which the channel material is mounted is P-type material and the gate is placed along the channel as shown in the figure. The thin layer of glass prevents any actual contact between the channel and the gate.

In operation, the source and the substrate are connected to the negative terminal of the battery and the drain is connected to the positive terminal. This will permit current to flow from the negative terminal of the battery to the source, through the channel to the drain and then back to the positive terminal of the battery.

When a negative voltage is applied to the gate, it has the effect of repelling electrons away from the gate as before. In addition, the negative potential applied to the gate attracts holes in the P-type material so that the width of the channel is reduced as shown in Fig. 37. Thus the current flow through the channel is restricted by the narrowing of the channel. In effect, the resistance of the channel is increased. We refer to this type of channel as a depletion channel. The transistor is called an insulated-gate-field-effect transistor and it is also referred to as a depletion type because the flow of current through the transistor is controlled by producing a depletion layer in the channel as in the case of the junction transistors discussed previously.

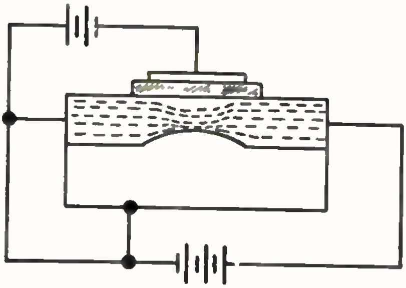


Fig. 37. Current flow through an N-channel IGFET with bias applied.

Both N-channel and P-channel IGFET's are manufactured. The schematic symbols used to represent the two different types are shown in Fig. 38A and B. In A, we have shown the symbol used for an N-channel type, and in B the schematic symbol used for a P-channel type. In operation, the units perform in essentially the same way as the junction-gate units with the exception that there will be no current

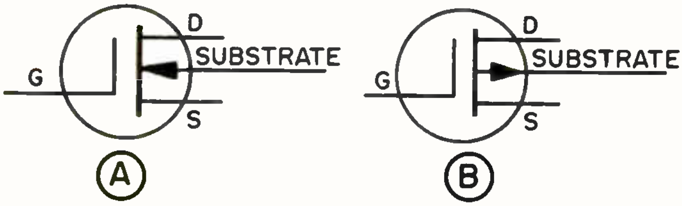


Fig. 38. Insulated-gate field-effect transistors. (A) shows the schematic symbol for an N-channel unit and (B) the symbol for a P-channel unit.

flow at all from the channel to the gate or from the gate to the channel. In the JFET, there may be very small leakage current across the junction. However, a JFET has a high input resistance because this leakage current is low. The IGFET has an even higher input resistance because there is no current flow at all from the gate to the channel or from the channel to the gate. Thus the input resistance of an IGFET is almost infinite.

Enhancement Type.

So far the field-effect transistors we have been discussing are all what are known as depletion types. In the depletion type of FET, the channel is formed and a bias is placed on the gate so as to reduce the size or width of the channel. In the enhancement-type of field-effect transistor, there is no channel present until the bias is applied to the gate. Thus, there is no current flow from the source to the drain through the transistor, unless there is a bias applied to the gate. The polarity of the bias applied to the gate is reversed from what it is in the depletion type, and this bias forms the channel through which current can flow. The operation of the units is the same as with the depletion type with the single exception of the reverse bias. In other words, in the case of an N-channel enhancement-type field-effect transistor, instead of placing a negative bias on the gate to reduce the width of the channel, as we do in the depletion-type transistor, in the enhancement-type we place a positive bias on the gate and produce the N channel.

The enhancement-type field-effect transistor is always an insulated gate type. In the case of a junction FET, if we produced an enhancement type, we would have current flow across the junction because the voltage required to produce the channel would forward bias the junction. However, in the insulated-gate FET, no current can flow across the junction because we have an insulating material between the gate and the channel. Thus we can put any type of bias we want, either forward or reverse bias, on the gate and we still will not get a current flow from the gate to the channel or from the channel to the gate.

The schematic symbol of an N-type IGFET of the enhancement type is shown in Fig. 39A. Notice that we have indicated there is no channel by breaking the channel into three parts. When the correct bias is applied to the gate, an N channel between the source and the drain will be formed. The schematic symbol for the P-channel unit of an enhancement-type IGFET is shown in Fig. 39B.

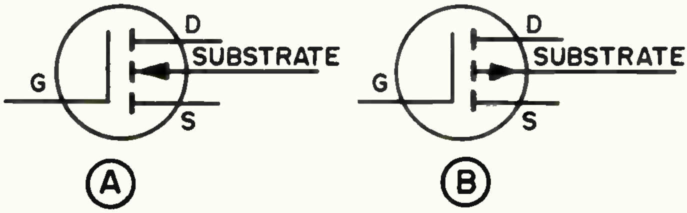


Fig. 39. A shows the schematic symbol for an N-channel enhancement-type IGFET. B shows the P-channel unit.

The operation of the enhancement-type IGFET is basically the same as with the depletion type. It could be used in a circuit similar to the circuits shown in Fig. 34 and Fig. 35.

One of the problems with IGFET's is the very high resistance between the gate and the channel. In shipping these units the manufacturer usually wraps the leads in tin foil to keep them connected together. If he doesn't do this, static charges can build up on the gate because of the very high resistance between the gate and the channel. These static charges may become high enough to actually puncture the insulation between the gate and the channel and thus ruin the unit.

In soldering an IGFET into a circuit, there might be enough leakage from the power line through the tip of your soldering iron to ruin the FET. To prevent this from happening, ground leads should be used on the various connections to the transistor and these leads should be left in place until the transistor is installed in the circuit. Once the transistor is soldered in place, you do not have to be concerned about static charges destroying the unit because the resistance in the circuit will be low enough to prevent static charges from building up to a high enough value to destroy the transistor.

Field-effect transistors are finding their way into commercial equipment, and you should therefore be sure you understand how they operate. You should review the sections on field-effect transistors several times if necessary because you can be sure they are going to be widely used in the future. They offer the advantages of the transistor as well as many of the advantages of the vacuum tube.

THE UNIJUNCTION

Another important semiconductor device is the unijunction. The unijunction is different from a conventional two-junction transistor in that it has only a single junction.

Most unijunctions are made of a bar of N-type silicon. There are two base contacts made to this bar called base 1 and base 2. These contacts are made at the ends of the bar. Between the two bases is a single rectifying contact called the emitter. The schematic symbol of the unijunction is shown in Fig. 40.

In Fig. 41 we have an equivalent circuit showing how the unijunction operates. We have referred to the resistance between base 1 and the emitter as R_{B1} and the resistance between base 2 and the emitter contact as R_{B2}. When a dc voltage is applied to the unijunction between B_1 and B_2, a current will flow through the base as shown. As long as the voltage drop across R_{B1} is greater than the emitter voltage, the emitter will be reverse biased so that there will be no current flow across the junction between the emitter and the base. The voltage across the resistance representing base 1 and the voltage across the resistance representing base 2 will remain constant. The two bases more or less act like two resistors

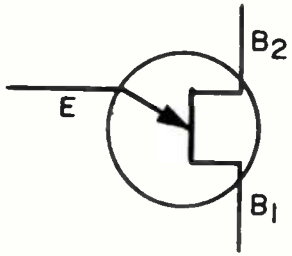


Fig. 40. Schematic symbol of a unijunction.

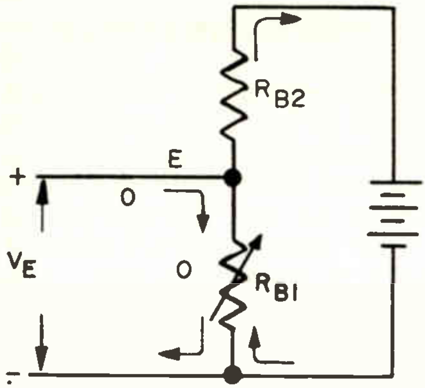


Fig. 41. Equivalent circuit showing the operation of the unijunction.

in series. The positive voltage at the emitter junction prevents any electrons from leaving the base and crossing the junction to the emitter and also prevents holes from traveling from the emitter to the base. There will be a small leakage current across the junction, but this is of no importance insofar as the operation of the unijunction is concerned.

If the voltage, V_E, exceeds the voltage across R_{B1}, then holes will enter the base and flow through R_{B1} as shown by the arrows on the diagram. These holes will cause the number of electrons flowing in R_{B1} to increase. The net result will be that you will have a drop in voltage across R_{B1} but at the same time an increase in current.

You will remember from Ohm's Law that the current flowing in a circuit is equal to the voltage divided by the resistance. If the voltage drops, the current must drop. However, in this device we have a situation where the voltage drops, but the current increases. We refer to this as "negative resistance". Devices that have this characteristic can be used in various types of amplifier circuits.

The unijunctions made for a number of years always made use of an N-type base material and a P-type emitter. However, recently some unijunctions using a P-type base material and an N-type emitter have been developed. The schematic symbol is the same, except that the direction of the arrow is reversed.

Unijunctions have not been widely used in commercial radio and TV equipment; however, they have been used in various pieces of test equipment. It is quite likely that as more transistorized television receivers are manufactured, the unijunction may be used in the sweep circuits since they are quite readily adapted to this type of application.

The important thing for you to remember at this time about the unijunction is that the device has a single junction and that the resistance of the two bases remains essentially constant until the emitter voltage exceeds the voltage across base 1. Then the voltage drop across base 1 decreases while the current flow through it increases, resulting in the negative resistance characteristic of base 1.

SUMMARY

There are too many details in this section to try to summarize them. The important thing for you to do is to realize that the different names assigned to the conventional two-junction transistors indicate the manufacturing process used to make the transistor. Typical two-junction transistors are either NPN or PNP transistors, and the basic theory of operation of the two-junction transistors is the same regardless of the manufacturing technique used. Different manufacturing techniques result in transistors with different characteristics, but the theory of operation is the same.

The field-effect transistor is a transistor that very closely resembles a vacuum tube in many of its characteristics. Remember that there are two basic types: the junction field-effect transistor and the insulated-gate field-effect transistor. In the insulated-gate type, the gate is insulated so that the leakage current to and from the gate is practically zero. This type of transistor has a very high input resistance.

You should also remember that field-effect transistors can be made in both N-channel types and P-channel types. You'll recall that by depletion type we are referring to a transistor where a channel is present. The input voltage to this type of transistor controls its channel width. JFET transistors are all of the depletion type. The IGFET may be either the depletion type or the enhancement type.

The unijunction is a semiconductor device with a single junction. Its use in commercial equipment is somewhat limited at this time, but you should understand the basic fundamentals of the device because it is quite likely that it will be used in the future.

One important point about all types of transistors that we must emphasize is that they are all easily damaged by excessive heat. This is true particularly of germanium transistors, but silicon transistors can also be destroyed by excessive heat. Whenever you have to replace a transistor in a circuit, you should make sure that the point at which you have to solder the transistor in the circuit is clean so that the solder will melt and flow over the connection quickly. Also make sure that the transistor leads are clean. It is a good idea to use a heat sink between the point at which you are soldering and the semiconductor device. A good heat sink is a pair of longnose pliers; simply hold the lead securely in the jaws of the pliers while you are soldering the lead in place. Much of the heat developed at the joint will flow through the pliers and keep the semiconductor device itself from becoming excessively hot. The joint should be soldered as quickly as possible; get the iron off the joint just as soon as the solder has melted and flowed smoothly over the connection.

Semiconductor devices can be damaged by storing them in excessively warm places. Again, this is particularly true of germanium transistors which are more heat sensitive than silicon transistors. Storing semiconductor devices at room temperature will prevent this type of damage. You should avoid storing them in any place where they can become excessively hot.

Now to check yourself on this important section you should answer the following self-test questions.

SELF-TEST QUESTIONS

(aw) Into what two basic types can the grown-junction transistor be divided?

(ax) What type of transistor can the surface-barrier transistor be classified as?

(ay) What is the most important characteristic of the surface-barrier transistor?

(az) What do we mean by a diffusion transistor?

(ba) What is an important use of the diffusion technique in manufacturing transistors?

(bb) What is the difference between a junction-gate field-effect transistor and an insulated-gate field-effect transistor?

(bc) What is a depletion-type field-effect transistor?

(bd) What is an enhancement-type FET?

(be) What is a unijunction?

Answers to Self-Test Questions

(a) Four.

(b) Germanium and silicon.

(c) A covalent bond is the sharing of two electrons by two atoms, one from each atom.

(d) Four. A single atom of germanium or silicon will share an electron from its outer ring and an electron from the outer ring of a nearby atom to form a covalent bond. It will do this with four electrons to establish four covalent bonds.

(e) Intrinsic conduction is conduction due to the formation of hole-electron pairs throughout a germanium or silicon crystal.

(f) No.

(g) Germanium.

(h) Heat.

(i) Silicon.

(j) An N-type material is a material that has been doped so that electrons are the majority carriers. This is brought about by using an impurity that has five electrons in the valence ring so that when it forms covalent bonds with nearby germanium or silicon atoms there will be an electron left over.

(k) A donor material is an impurity which when added to silicon or germanium will form covalent bonds with four nearby atoms and have an electron left over. When a donor material is added to germanium or silicon, N-type material is formed.

(l) Arsenic, antimony, and phosphorous.

(m) P-type semiconductor material is a material that has been doped with an impurity having three electrons in the valence ring. This will leave a covalent bond that is short one electron so there will be a hole in the bond. The hole is in effect a positive charge and hence the majority carriers in the P-type material are the holes or positive charges.

(n) An accepter-type impurity is an impurity with three electrons in the valence ring or shell. It is an accepter-type material because it leaves a hole in the covalent bond which can accept an electron.

(o) Indium, boron and aluminum.

(p) Electrons are the majority carriers in N-type material.

(q) Holes.

(r) When the arsenic loses an electron it will be short one electron to completely neutralize the charge on the nucleus, and therefore the atom will have a positive charge. Meanwhile the atom of silicon or germanium that has received the extra electron will have a negative charge on it.

(s) There is no charge on the crystal, it is neutral. Although some regions may have a positive charge, other regions may have a negative charge; the crystal itself neither gains nor loses electrons and therefore it does not have any charge.

(t) Holes are produced.

(u) Diffusion is a random motion of the carriers in a semiconductor material. It goes on at all times in the crystal and every effort is made to keep diffusion as low as possible since it contributes nothing insofar as the usefulness of the material in semiconductor devices is concerned.

(v) Drift.

(w) Electrons are the majority carriers in an N-type material and they move from the end to which the negative potential is applied towards the end to which the positive potential is applied.

(x) Holes are the majority carriers in a P-type material and they move from the end to which the positive potential is applied to the end to which the negative potential is applied.

(y) No - the crystal will remain electrically neutral. In the case of N-type material, exactly the same number of electrons will leave the positive end of the crystal and enter the negative end of the crystal. In the case of the P-type material, electrons will leave the end to which the positive potential is connected creating holes. Exactly the same number of electrons will enter the end to which the negative potential is connected to fill holes arriving at the negative end.

(z) No. For a given potential and given size of crystal, electrons will move at approximately twice the rate through an N-type crystal as the holes will through a P-type crystal.

(aa) The N-type material will have the lower resistance. This is due to the higher mobility of the electrons in the N-type material than the holes in the P-type material.

(ab) Detectors and rectifiers.

(ac) The depletion layer is an area on both sides of the junction. On the P-side of the junction there is a shortage of holes and on the N-side of the junction there is a shortage of electrons. The shortage is caused by a few of the majority carriers crossing the junction in each way building up a charge at the junction so that the majority carriers are repelled away from the junction.

(ad) The potential barrier is the voltage built up across the junction by the diffusion of majority carriers across the junction. The holes that diffuse across the junction into the N-side of the junction create an area that has a negative charge in the P-side of the junction. Similarly, the electrons diffusing across the junction into the P-side create an area on the N-side of the junction that has a positive charge. This charge across the junction eventually becomes high enough to prevent any further diffusion of holes and electrons across the junction.

(ae) No. The net charge on the crystal will remain zero. There may be areas on the crystal that have a positive charge, and other areas that have a negative charge, but since the crystal itself neither gains nor loses electrons, the net charge on the crystal will remain zero.

(af) Yes. Minority carriers crossing the junction tend to weaken the potential barrier established across the junction by majority carriers diffusing across the junction. When the potential barrier is weakened, additional majority carriers can cross the junction. Thus we end up with carriers crossing the junction in both directions. This adds nothing to the useful current that the diode can handle, but it does contribute to heating and thus limits the useful current that can cross the junction.

(ag) When a junction is forward biased we have a positive potential applied to the P-side and a negative potential applied to the N-side. This permits electrons to freely cross the junction from the N region to the P region. Similarly holes can cross the junction from the P region to the N region.

(ah) When a junction is reverse biased we have a negative potential connected to the P region and a positive potential connected to the N region. The positive potential connected to the N region repels holes in the P region away from the junction so that they cannot cross the junction. Similarly, the negative potential applied to the P region repels electrons in the N region away from the junction so that they cannot cross the junction. When a junction is reverse biased, majority carriers normally cannot cross the junction.

(ai) When there is no voltage applied to a semiconductor diode, the net current flow across the junction is zero. However, in the case of a vacuum tube where there is no voltage applied between plate and cathode, some electrons will leave the cathode with sufficient energy to travel over to the plate. As a result, there will be a small current through the tube even though there is no voltage applied between the plate and cathode.

(aj) When a semiconductor diode is reverse biased, there will be a small current flow across the junction due to minority carriers. As long as the breakdown voltage of the diode is not exceeded, this current will be quite small. In the case of a vacuum tube, when the plate is made negative with respect to the cathode, the plate will repel electrons so that there will be no current flow through the vacuum tube.

(ak) A Zener diode is a diode used in applications where a reverse bias is placed across the junction. The diode is designed to break down at a certain voltage and then maintain a constant voltage. If the voltage tries to increase above this constant value, the current flow through the Zener diode will increase so that the diode can be used in voltage regulating circuits and also can be used as a voltage reference source.

(al) A tunnel diode is a diode where the electrons cross the junction by a process similar to tunneling across the junction. The tunnel diode has a characteristic of introducing negative resistance into the circuit when a certain voltage is applied across the junction. In other words, when the voltage across the diode increases, the current flow through the diode decreases. Similarly, when the voltage decreases the current increases. Because of this negative resistance characteristic, the tunnel diode can be used as an oscillator.

(am) A p-i-n diode is a diode that is primarily used as a variable resistance. The resistance of the diode varies as the voltage across it is varied. The p-i-n diode is used in automatic gain control circuits to vary the strength of the signal reaching amplifier stages.

(an) The base region is the center region of the transistor. On one side of the base region is the emitter, and on the other side is the collector.

(ao) Germanium and silicon.

(ap) PNP transistors and NPN transistors.

(aq) Forward bias.

(ar) Reverse bias.

(as) The base region is usually comparatively thin.

(at) See Fig. 22.

(au) See Fig. 24.

(av) Holes are useful current carriers in a PNP transistor.

(aw) NPN and PNP transistors.

(ax) An alloy-type transistor.

(ay) Good high-frequency performance.

(az) A diffusion transistor is a transistor which has been made by diffusing the impurities into the emitter, base and collector regions.

(ba) One of the most important uses of the diffusion technique is in the manufacture of non-uniform base regions.

(bb) In a junction-gate field-effect transistor there is an actual contact between the channel material and the gate. There will be some current flow across the contact at all times due to minority carriers crossing the junction. In addition, if the junction is forward biased there will be a high current flow across the junction. In an insulated-gate field-effect transistor a glass or similar insulating material is used between the material in the channel and the gate. Since there is an insulator between the gate and the channel, there will be little or no current flow across the insulator either due to minority carriers when there is a reverse bias applied, or due to majority carriers with a forward bias applied.

(bc) A depletion-type FET is a unit in which the channel is present at all times. The transistor works by depleting or reducing the size of the channel.

(bd) An enhancement FET is a unit in which there is no channel present until the operating bias is applied between the gate and the material in which the channel is formed.

(be) A unijunction is a semiconductor device having two base connections but only a single junction. The junction is called the emitter. The single junction makes the unijunction quite different from the conventional two-junction transistor.

Lesson Questions

Be sure to number your Answer Sheet B112.

Place your Student Number on every Answer Sheet.

Most students want to know their grades as soon as possible, so they mail in their answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable. However, don't hold your answers too long; you may lose them. Don't hold answers to more than two sets of lessons at any time, or you may run out of lessons before new ones can reach you.

1. Name the two most important semiconductor materials used for transistors.

2. When a donor type of material is added to a silicon or a germanium crystal, what type of semiconductor material is produced? Does this type have free electrons or free holes?

3. What effect do the two layers of ionized atoms at the junction in a PN diode have on the majority carriers in the vicinity of the junction?

4. To which side of a PN junction diode do you connect the positive battery terminal if you wish to place a forward bias on the junction?

5. If a reverse bias is applied to a junction diode, what effect will a small increase in bias have on the current flowing, provided the reverse voltage does not exceed the breakdown voltage?

6. What is a Zener diode?

7. In a PNP transistor, what happens to a hole that crosses the emitter and the base and moves into the collector?

8. What is a drift transistor?

9. What is an N-channel, junction-type field-effect transistor?

10. What do we mean when we refer to a field-effect transistor as an enhancement type?

CASHING IN ON DISCONTENT

Discontent is a good thing--if it makes you want to do something worthwhile. If you had not been discontented, you would never have enrolled for the NRI course.

Practically everyone is discontented. But some of us are "floored" by discontent. We develop into complainers. We find fault with anything and everything. We end up as sour and dismal failures.

Those of us who are wise use our discontent as fuel for endeavor. We keep striving toward a goal we have set for ourselves. We are happy in our work. We face defeat, and we come out the victors.

At this minute you may be discontented with many things--your progress with your course, your earning ability, yourself.

Make that discontent pay you dividends. Don't let it throw you down. If you do, you may never be able to get up again. Keep striving to remove the cause of your discontent. Remember that it's always darkest before the dawn. And a real NRI man works hardest and accomplishes most when he is face to face with the greatest discouragements.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-212

SERVICING • ELECTRONICS • COMMUNICATIONS

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

HOW TRANSISTORS WORK

B112

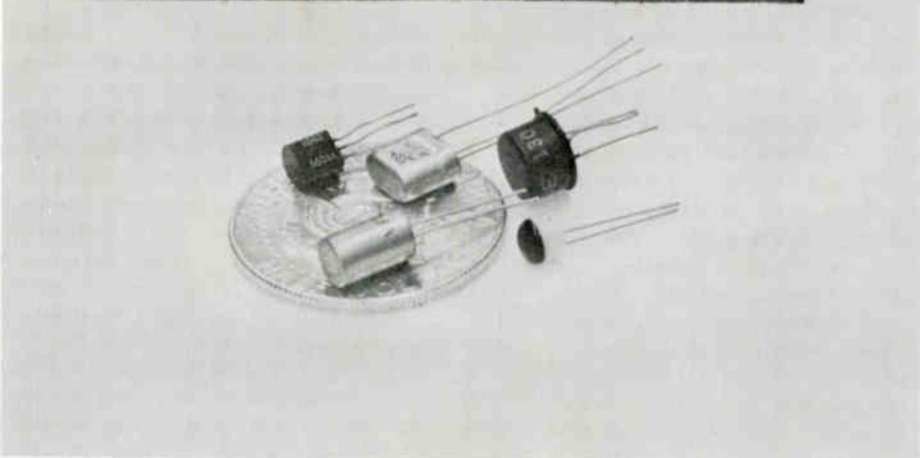
HOW TRANSISTORS WORK

B112

STUDY SCHEDULE

JD4M168 **1968 EDITION** Printed in U.S.A.

HOW TRANSISTORS WORK



In the preceding lessons you studied tubes, and you saw how they are used in different circuits. In this lesson you will study semiconductor devices - these devices have already replaced tubes in many important applications and are rapidly moving into new areas that were once dominated entirely by tubes.

An example of the importance of semiconductor devices can be seen in entertainment-type equipment. Just a few years ago all the rectifiers used in this equipment were vacuum tubes. Today, however, the vacuum tube is no longer used for this purpose; rectifiers in the modern entertainment-type equipment are all semiconductor devices.

Semiconductor devices used as rectifiers have two elements and are called diodes just as two-element vacuum tubes are called diodes. Semiconductors used to amplify signals usually have three or more elements and are called transistors. There are a large number of different types of transistors available today, but for the most part these transistors can be classified into two types, the NPN transistor and the PNP transistor. If you understand how these two transistor types work, you should have little difficulty understanding how all others work and any new transistors that might be introduced in the future. You will run into many different types of transistors identified by different names, but these names usually refer to the method used in manufacturing the transistor rather than the manner in which it operates.

There are some similarities between tubes and semiconductors. A two-element vacuum tube can be used to change an alternating current to a direct current; a two-element semiconductor can be used for the same purpose. A triode vacuum tube can be used to amplify a signal; a transistor can be used for the same purpose. However, this is where the similarity ends. Most tubes are vacuum devices; in other words, all the air and gas have been evacuated from inside the tube. On the other hand, a semiconductor is a solid device and there is no space between the elements in it. We have a current flow through a vacuum in a tube, but we have a current flow through a solid in a semiconductor.

The importance of semiconductors cannot be overemphasized. They have completely supplanted the vacuum tube in portable radio receivers and in automobile receivers. Almost all high fidelity and stereo' equipment manufactured today uses semiconductors exclusively - the only tube-operated equipment of this type you are likely to encounter is equipment that is several years old. Semiconductors are finding their way into television receivers and it is probably just a matter of time before they completely replace the vacuum tube.

Semiconductors have several advantages over the vacuum tube. Perhaps one of the most important advantages is that they do not require any heater or filament power. Not only is this a power saving in the operation of the equipment, but it also removes considerable heat from the equipment. Heat is probably the thing that causes the most damage to parts in electronic equipment. Thus with the removal of the heater or filament power from the equipment, other components such as capacitors, etc. will last longer.

Semiconductors are very rugged. They are solid devices and hence not subject to breakage from mechanical shock as tubes are. An important advantage of transistors is that they will operate on a comparatively low voltage, and this usually results in some reduction of the power required in the equipment.

Although semiconductors have many advantages over vacuum tubes, they do have some disadvantages. One disadvantage is that it is usually not possible to get as high a gain in an amplifier stage using a transistor as it is in a similar stage using a tube. Therefore to get the equivalent gain, more transistor stages are required than vacuum-tube stages. Another disadvantage of the transistor is that its characteristics are not as constant as those of a vacuum tube. In other words, you are more likely to run into difficulty replacing a transistor than you are in replacing a tube because the replacement transistor's characteristics might be considerably different from the characteristics of the original transistor. Another disadvantage of both diode semiconductors and transistors is that their characteristics can vary appreciably with changes in temperature. As a matter of fact, some semiconductor devices are easily destroyed by too much heat.

In spite of the fact that semiconductor devices have some disadvantages when compared to vacuum tubes, their advantages more than outweigh the disadvantages and their importance in the field of electronics is continually growing. Therefore it is important that the technician have a good understanding of semiconductor fundamentals, how they are used, and how they operate. Before going ahead to see how semiconductors are used as rectifiers and amplifiers, we need to know more about certain types of atoms, in order to understand how these devices work.

Semiconductor Fundamentals

You have already learned that certain materials will conduct electricity readily and that some materials will hardly pass any electric current at all. The materials that will conduct current readily are called conductors and those that will not conduct current are called insulators. Midway between the two types of materials is a group of materials called semiconductors. These materials are not good conductors, nor are they particularly good insulators. Two examples of semiconductor materials are germanium and silicon. These are the materials that we will be mostly concerned with in this section. Both diode semiconductors and transistors are made from germanium and silicon. A new material that shows promise for use in semiconductors is gallium arsenide. It's likely that this material will be used in semiconductors in the future. Before going ahead with our detailed study of semiconductor materials, let us review a few important facts about conductors and insulators.

CONDUCTORS AND INSULATORS

You will remember that all materials are made up of atoms. An atom is the smallest particle of a material that retains the characteristics of the material.

In the center of the atom is the nucleus. This nucleus contains a positive charge. The number of positive charges on the nucleus distinguishes one material from another. In other words, the nucleus of a copper atom does not have the same number of positive charges as the nucleus of an iron atom.

Each atom normally has enough electrons, which have a negative charge, to exactly neutralize the positive charge on the nucleus. Thus, the hydrogen atom which has a nucleus with one positive charge will have one electron, and the helium atom which has a positive charge of two in the nucleus will have two electrons. Another atom that has a nucleus with 30 positive charges will have 30 electrons to exactly neutralize the positive charge on the nucleus.

The electrons in an atom arrange themselves in shells around the nucleus. The total number of electrons will normally be just enough to neutralize the charge on the nucleus. However, there is a maximum number of electrons that can be forced into each shell. In the first shell around the nucleus, the maximum number of electrons is 2. In the second shell, the maximum number of electrons is 8, and in the next shell the maximum number of electrons is 18. A shell can have less than the maximum number of electrons, but not more than the maximum number.

Conductors.

An example of an atom in a conductor is shown in Fig. 1. We have drawn the shells in the form of rings, but remember that this atom actually has three dimensions, not two. Notice that in this atom there are two electrons in the first shell,

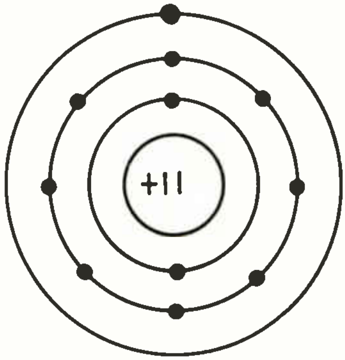


Fig. 1. An atom of a conductor.

8 electrons in the next shell and only one electron in the third shell. The outer shell is called a valence shell. The single electron in the third shell, which is called the valence electron, is not very closely bound to the nucleus; it can easily be removed from the atom. Thus a material of this type has a large number of electrons that can easily be removed from their atoms. When these electrons are forced to move in one direction we have a current flow. Thus a material that has only one or two electrons in an outer shell that could have many more, is a conductor, because the one or two electrons in the outer shell are not closely bound to the nucleus.

Insulators.

An atom of an insulator is shown in Fig. 2. Notice that in this atom

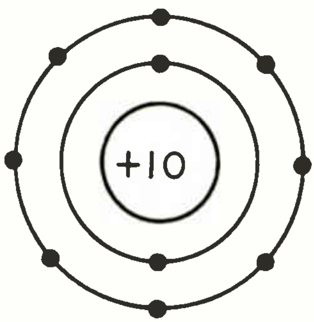


Fig. 2. An atom of an insulator.

there are two electrons in the first shell, and 8 electrons in the second shell. Both the shells are completely filled and will be closely bound to the nucleus. This means that it is very difficult to get one of these electrons out of an atom and therefore this material is an insulator or nonconductor.

Remember the important difference between conductors and insulators. A conductor is a material that has one or two electrons in the outer shell that are not closely bound to the nucleus, whereas an insulator is a material in which the outer shell of each atom is filled or almost filled so that the electrons are closely bound to the atom and cannot be easily removed. Because these electrons cannot be removed from the atom, this type of material normally will not conduct current, and hence is called an insulator.

SEMICONDUCTOR MATERIAL

A material that is classified as a semiconductor has electrical characteristics midway between those of a conductor and those of an insulator. The electrons in a semiconductor can be removed from their atoms when some type of external energy, such as voltage, heat, or light is applied to the material. Then the material acts like a conductor.

The most important semiconductor materials used for transistors are germanium and silicon. The first low-cost transistors were germanium transistors, but recent developments have lowered the cost of silicon transistors so that most of the new transistor types being introduced are silicon. Both germanium and silicon are very abundant elements, but neither is found in the pure state, and it is quite difficult to process them to the high state of purity required for use in transistors. The first transistors were made of germanium because techniques for getting pure germanium were developed first. However, now it is possible to refine silicon to the high degree of purity required, at a reasonable cost, and since silicon has several advantages over germanium for use in semiconductor devices it has in many ways replaced germanium.

In general, there is not too much difference between the operation of semiconductor devices made from germanium and those made from silicon. We will cover the important points of these devices made from both materials since you will run into semiconductor devices of both types.

The Germanium Atom.

The arrangement of electrons about the nucleus in a germanium atom is shown in Fig. 3A. The nucleus of the germanium atom has a positive charge of 32. Thus as you might expect, there will be 32 electrons revolving in the shells about the nucleus. There are two electrons in the first shell, eight in the second, eighteen in the third and four in the fourth shell. Thus, the first, second and third shells are filled, but there are only four electrons in the outer shell. However, these four electrons, which are called the valence electrons, are bound to the nucleus much more so than the one or two electrons found in the outer shell of a conductor.

The important electrons in the germanium atom, insofar as its use in semiconductors is concerned, are

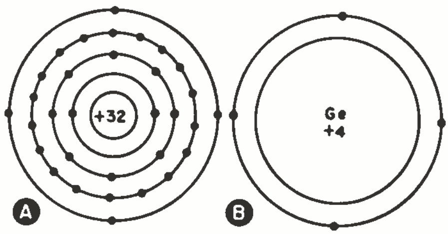


Fig. 3. (A) is the germanium atom with a charge of 32. (B) shows the simplified symbol.

the four electrons in the outer shell because the shell is not filled. The other electrons are bound so closely to the nucleus that they cannot easily be removed. Therefore, germanium is often represented as shown in Fig. 3B.

The Silicon Atom.

The arrangement of electrons about the nucleus in a silicon atom is shown in Fig. 4A. The nucleus of the silicon atom has a positive charge of 14. Therefore, there will be fourteen electrons revolving about the nucleus. There are two electrons in the first ring, eight in the second and four in the third. Thus the first and second rings are filled, but there are only four electrons in the outer shell. These four

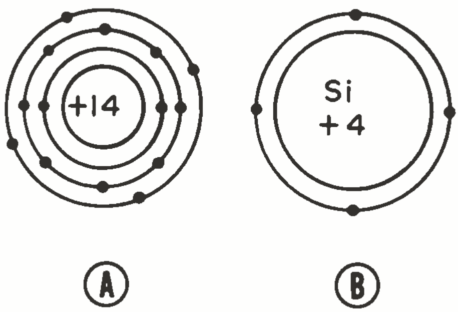


Fig. 4. (A) shows the silicon atom with a charge of 14. (B) shows the outer shell with four electrons.

electrons are the valence electrons like the four in the germanium atom and are bound fairly closely to the nucleus. As in the case of the germanium atom, the four electrons in the outer ring are the ones that are of importance in the use of silicon in semiconductors.

Notice the similarity between the silicon and germanium atoms. In both atoms, the outer shell or ring has four electrons, and all the other shells are filled.

The tendency of some materials like silicon and germanium, that do not have the outer shell completely filled with electrons, is to get additional electrons to fill up the outer shell. In pure germanium and silicon, the electrons in the outer shell of one atom are bound as closely to that atom as the four electrons in the outer shell of another atom. Therefore one atom cannot pull electrons away from another atom. Instead, two nearby atoms will share one outer electron from each atom. In other words, two atoms of germanium may share electrons as shown in Fig. 5A; and two atoms of silicon may share electrons as shown in Fig. 5B. By sharing electrons in this way, each atom will partly fill its outer shell. This pair of shared electrons, one from each of two atoms, is called "a covalent bond".

In order to try to fill its outer ring with electrons, a single germanium atom or a single silicon atom will establish covalent bonds with four other atoms. This arrangement of atoms in a piece of germanium is shown in Fig. 6A. A similar arrangement of atoms in a piece of silicon is shown in Fig. 6B. These pieces of silicon and germanium are called crystals and the way in which they are arranged is called a lattice structure. Each atom shares each of its four valence electrons with one valence electron of another atom to form these bonds.

INTRINSIC CONDUCTION

Even at comparatively low temperatures, there is heat energy in all materials. This energy is sufficient to cause a few of the electrons to move out of their proper place in the lattice structure of either the germanium crystal or the silicon crystal and become free electrons. These free electrons are available for conduction of electric current. The number of free electrons available is much higher in germanium than it is in silicon.

When one of these electrons moves out of its position in the lattice structure, it leaves an empty space in the crystal lattice. This empty space is called a hole. An electron from a

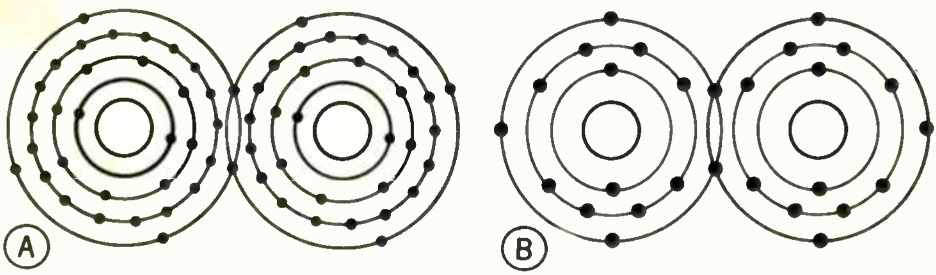


Fig. 5. The sharing of two electrons by two germanium atoms is shown at A; by two silicon atoms at B.

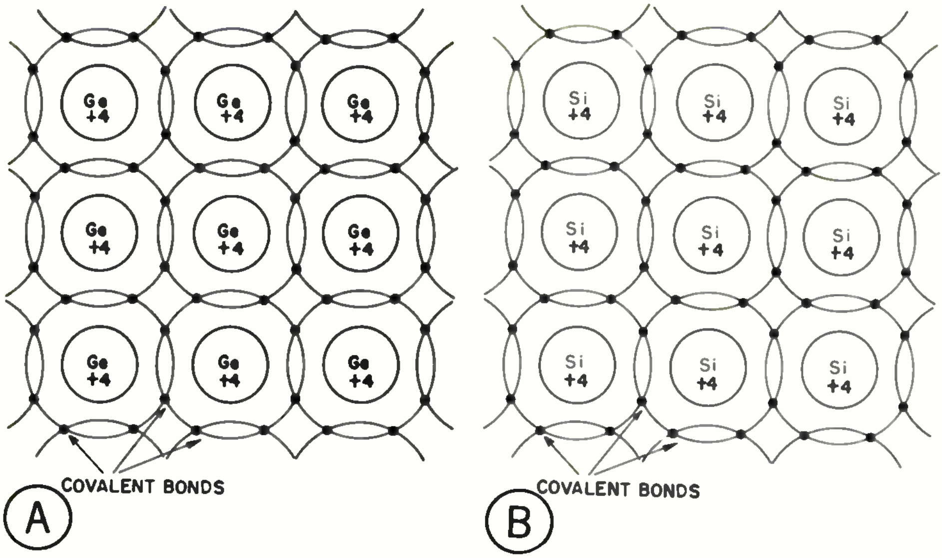


Fig. 6. The lattice structure of germanium is shown at A; of silicon at B.

nearby atom can move into this hole thus creating a new hole at the place it left. Another electron may move out of still another atom to fill this new hole, leaving behind it a hole. This movement of an electron to fill a hole thus creating a new hole in the place it left makes it look as if holes themselves move. Furthermore, since the hole represents a missing electron, it has a positive charge.

In a piece of germanium or silicon, the electrons are in a constant state of motion about their atoms. If in its movement an electron comes closer to a hole than to its own atomic nucleus, it will be strongly attracted to the hole and will leave its atom. When there is no voltage applied across the crystal, the movement of a hole or an electron is a random movement. Holes and electrons may move in any direction.

If heat or some other form of external energy is applied to the crystal, the resistance of the material is reduced. This happens because more electrons are freed by the energy applied to the crystal. In addition, the speed of the random movement is increased.

The movement of an electron out of an atom forms a hole in the atom. Thus, whenever an electron is freed from an atom a hole is formed. This free electron and the hole it forms are called a "hole-electron pair". The formation of hole-electron pairs is a continuous process. Also the filling of holes by electrons is a continuous process. In other words, the process of an electron leaving its atom and forming a hole, and another electron moving in to fill the hole and in so doing creating a new hole, is a continuous process. The conduction of electricity in pure germanium or pure silicon crystals due to the formation of hole-electron pairs is called the intrinsic conduction.

The conductivity of a germanium crystal or a silicon crystal, which is the ability of the material to conduct an electric current, depends on the average length of time an electron is free and on the number of free electrons. We mentioned previously that there are more electrons free if external energy, such as heat, is applied to the material. Therefore the conductivity rises as the temperature of the material is increased.

This type of conduction is much higher in germanium than it is in silicon. As an example, if we had a germanium crystal exactly one centimeter on each side and measured the resistance across two parallel surfaces, we would find the resistance to be approximately 60 ohms. The resistance of an equivalent piece of silicon would be approximately 60,000 ohms. Thus intrinsic conduction is much higher in germanium than it is in silicon.

Intrinsic conduction in transistors is undesirable. It is kept as low as possible by holding the operating temperature of the material down. Transistors are also shielded from light because light is a form of energy and light striking the crystal will increase the intrinsic conduction. Since silicon has a much lower intrinsic conduction than germanium, semiconductors made from silicon are less affected by heat than are semiconductors made from germanium. This is one of the chief advantages of silicon over germanium as a semiconductor material.

In their pure forms, neither germanium nor silicon are useful in semiconductor devices. In fact, in spite of intrinsic conduction, neither material is a good conductor at room temperature; they are both fairly good insulators. To use these materials in semiconductors, controlled amounts of other selected elements called impurities are added to the crystals to alter their characteristics. By adding these materials we can produce two types of silicon and two types of germanium. They are called N-type and P-type. Now, let us study the characteristics of these two types of materials.

N-TYPE MATERIAL

N-type silicon or germanium can be produced by adding as an impurity an element that has five electrons in its outer ring. An example of this type of material is arsenic. Arsenic has a positive charge of 33 on the nucleus and has 33 electrons in the shells surrounding the nucleus. There are two electrons in the first shell, eight in the second, eighteen in the third and five in the fourth or outer shell. In other words, arsenic is just like germanium except that the nucleus has one more positive charge and there is one additional electron in the outer shell.

If a small amount of arsenic is added to the germanium, the arsenic atoms will form covalent bonds with the germanium atoms as shown in Fig. 7. However, to form the covalent bonds with its neighboring germanium atoms, the arsenic atom needs only four of the electrons in its outer shell. Therefore there will be one electron left over when the arsenic atom forms covalent bonds with the four neighboring germanium atoms. This electron is free to move about within the crystal in exactly the same manner as a single valence electron in a good metal conductor. The addition of arsenic, which produces these free electrons, greatly reduces the resistance of the material.

When a small amount of arsenic

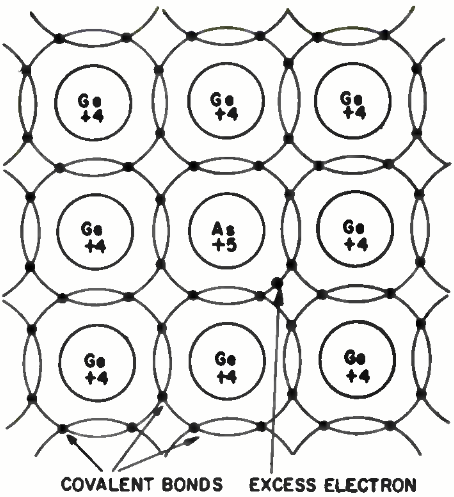


Fig. 7. Germanium with arsenic added.

is added to a silicon crystal exactly the same thing happens. The arsenic atom forms covalent bonds with the silicon atoms. As in the case of the germanium atom, only four of the electrons in the outer shell of the arsenic atom are used in forming these covalent bonds so there will be one electron left over.

When germanium or silicon have had an impurity added to them we say they have been doped. When semiconductor material has been doped with a material such as arsenic that results in there being excess electrons, we call it an N-type material. The N refers to the negative carriers, which are the free electrons. Arsenic is called a donor impurity because it donates an easily freed electron.

In addition to arsenic, other materials have been used as donors. Phosphorus, which has a total of fifteen electrons, can be used. The phosphorus atom has two electrons in the first shell, eight in the second and five in the third. Four of the electrons in the valence shell or ring will form covalent bonds with germanium or silicon atoms leaving a fifth electron free. Antimony, which has 51 electrons, also has been used as a donor. Antimony has two electrons in the first shell, eight in the second, eighteen in the third, eighteen in the fourth and five in the fifth or valence shell.

P-TYPE MATERIAL

If instead of adding a material with five electrons in its valence shell, we add a material with only three electrons in the valence shell, we have a situation where the impurity added to the silicon or germanium has one less electron than it needs to establish covalent bonds with four neighboring atoms. Thus, in one covalent bond there will be only one electron instead of two. This will leave a hole in that covalent bond.

A material that is frequently used for this purpose is indium. Indium has 49 electrons, two in the first shell, eight in the second, eighteen in the third, eighteen in the fourth and three in the fifth or valence shell.

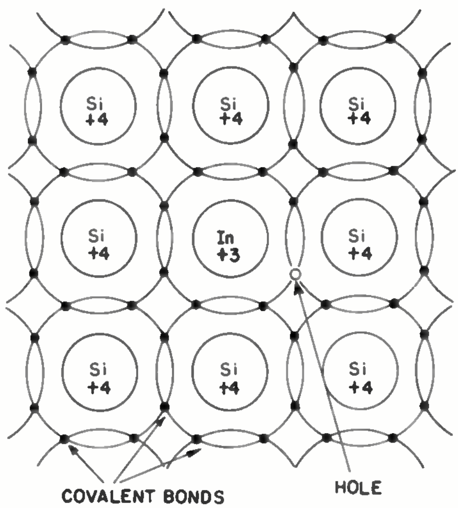


Fig. 8. Silicon with indium added.

The manner in which indium forms covalent bonds with neighboring silicon atoms is shown in Fig. 8. It forms covalent bonds with germanium atoms in the same way.

We mentioned previously that even at comparatively cold temperatures there is some heat energy within the crystal and thus there will be a few free electrons moving about the crystal. These free electrons are strongly attracted to the holes in the covalent bond produced where an indium atom has replaced a silicon or a germanium atom. Thus an electron will move into a hole in the covalent bond producing a new hole in another atom and giving the effect that the hole is moving as shown in Fig. 9.

Since a hole in the crystal actually represents a shortage of an electron, it is an area with a positive charge. Therefore when a semiconductor material has been doped with a material such as indium that produces holes in the lattice structure, we call it a P-type material.

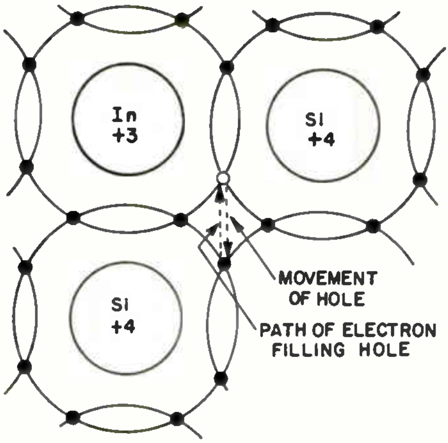


Fig. 9. When an electron fills a hole, another hole will apparently move to where the electron was.

P stands for positive; since holes represent a shortage of an electron we say they act as positive carriers. The indium is called an accepter impurity because its atoms leave holes in the crystal structure that are free to accept electrons. In addition to indium, boron and aluminum are also used as accepter impurities. Boron has 5 electrons, two in the first shell and three in the second which is the valence shell. Aluminum has 13 electrons, two in the first shell, eight in the second and three in the third or valence shell.

CHARGES IN N-TYPE AND P-TYPE MATERIAL

When a donor material such as arsenic is added to germanium or silicon, the fifth electron in the valence ring of the arsenic atom does not become part of a covalent bond. This extra electron may move away from the arsenic atom to one of the nearby germanium or silicon atoms.

The arsenic atom has a charge of +33 on the nucleus and normally has 33 electrons to neutralize this charge. When the electron moves away from the atom there will be only 32 electrons to neutralize the charge on the nucleus, and as a result there will be a small region of positive charge around the arsenic atom. Similarly, the excess electron that has moved into a nearby germanium or silicon atom will provide an excess electron in the atom. In the case of the germanium atom there will be a total of 33 electrons around a nucleus requiring only 32 electrons to completely neutralize it, and in the case of the silicon atom there will be 15 electrons around the nucleus requiring only 14 electrons to neutralize it. This means that the atom will have an extra electron so that there will be a region of negative charge around this atom.

It is important to notice that although there is a region of positive charge around the arsenic atom after the electron has moved away, and a region of negative charge around the germanium or silicon atom taking up the extra electron, the total charge on the crystal remains the same. In other words, a given crystal will have a net charge of zero. This means that there will be exactly enough electrons to neutralize the positive charges on the nuclei on the various atoms. But because some of the electrons may move about in the crystal, there will be regions in the crystal where there are negative charges and other regions where there are positive charges, even though the net charge on the crystal is zero.

In a P-type material to which material such as indium has been added we will have a similar situation. You will remember that the indium atom has only three electrons in its valence ring. These are all that were needed to neutralize the positive charge on the nucleus. However, with only three electrons in the valence ring, there is a hole in one of the covalent bonds formed between the indium atom and the four adjacent germanium or silicon atoms. If an electron moves in to fill this hole, then there is one more electron in the indium atom than is needed to neutralize the charge on the nucleus. Thus there will be a region of negative charge around the indium atom. Similarly, if one of the germanium or silicon atoms has given up an electron to fill the hole in the covalent bond, then the atom which has given up the electron will be short an electron so that there will be a region of positive charge around this atom. Again, while this giving up of an electron by a germanium atom and the acceptance of an electron by the indium atom ionizes or charges both atoms involved, the net charge on the crystal is still zero. We simply have one atom that is short an electron and another atom that has one too many. The crystal itself does not take on any charge.

These ionized atoms produced in both the N-type and the P-type germanium and the N-type and the P-type silicon are not concentrated in any one part of the crystal, but instead are spread uniformly about the crystal. If any region within the crystal were to have a very large number of positively charged atoms, these atoms would attract free electrons from other parts of the crystal to neutralize part of the charged atom, so that the charge would be spread uniformly about the crystal. Similarly, if a large number of atoms within a small region have had an excess of electrons, these electrons would repel each other and spread throughout the crystal.

Both holes and electrons are involved in conduction at all times. Holes are called positive carriers and electrons negative carriers. The one present in greatest quantity is called the majority carrier; the other is the minority carrier. In an N-type material, electrons are the majority carriers and holes are the minority carriers, whereas in a P-type material, holes are the majority carriers and electrons the minority carriers.

SUMMARY

This is a very important section of this lesson. You have covered many of the fundamentals of semiconductors on which we will build the remainder of the lesson. It is important that you understand the basic theory of semiconductors in order to be able to understand how semiconductor diodes and transistors work. We will summarize the important points that were covered in the preceding section.

If any of these points are not clear, you should go back and study the lesson again until they are clear. If you understand the first section of the lesson, you should be able to understand the material following without too much difficulty. However, if you do not understand what has been covered previously, you will have difficulty understanding what is to follow.

Pure semiconductor material such as germanium or silicon is a very poor conductor. In fact, it is an insulator if it is protected from all outside sources of energy. However, even at room temperature there is enough heat present in germanium and silicon to produce some electron and hole movement. The movement is much greater in germanium than it is in silicon.

An electron movement out of a covalent bond in a germanium or silicon atom leaves a hole in that bond. The hole will attract an electron from a nearby atom, producing a hole in that atom. Thus, both the hole and the electron appear to move. The holes are positive carriers and the electrons negative carriers of electricity. This formation of hole-electron pairs is undesirable in transistors and steps are taken to keep it as low as possible. The formation of hole-electron pairs increases as the temperature increases and is a much more serious problem in germanium-type semiconductor material than in silicon-type semiconductor material.

Semiconductor materials can be doped by adding small amounts of impurities. If a material with five electrons in the valence ring is added, the material is called a donor-type impurity. This type of material has one electron left over after it forms covalent bonds with four neighboring germanium or silicon atoms. Thus there will be an excess of electrons. We then refer to this kind of material as an N-type material.

If the germanium or silicon is doped with an impurity, called an accepter impurity, having three electrons in the valence shell, the impurity forms covalent bonds with the four neighboring germanium or silicon atoms. However, there will be a hole in one of the covalent bonds because the impurity has only three electrons available to form covalent bonds with four neighboring germanium and silicon atoms. This type of germanium or silicon is called P-type because there will be holes in the material, and these holes act as positive carriers.

Another point to remember is that when an electron is freed or when a hole captures an electron, the atoms involved become charged, or ionized. Thus throughout both N-type and P-type germanium or N-type and P-type silicon we have small regions of charge. However, the net charge on the crystal is zero and the charged regions are evenly distributed throughout the material.

SELF-TEST QUESTIONS

(a) How many electrons are there in a valence shell or ring in a silicon atom?

(b) What are the two types of material most widely used in semiconductor devices?

(c) What is meant by a covalent bond?

(d) How many covalent bonds will a single germanium or silicon atom establish?

(e) What is intrinsic conduction?

(f) Is intrinsic conduction desirable?

(g) In which type of conductor material, germanium or silicon, is intrinsic conduction the greatest?

(h) What is the greatest cause of an increase in intrinsic conduction in germanium?

(i) Which semiconductor material, silicon or germanium, has greater resistance?

(j) What is an N-type material?

(k) What is a donor material?

(l) Name two materials used as donors.

(m) What is P-type material?

(n) What is an accepter impurity?

(o) Name three types of accepter material.

(p) In N-type material what is the majority carrier?

(q) What are the majority carriers in the P-type material?

(r) When a donor impurity such as arsenic loses an electron in a semiconductor material, what happens to the arsenic and to a nearby atom that gains the electron insofar as their relative charge is concerned?

(s) Although there are small areas that have positive and negative charges in a doped semiconductor, what is the overall charge on the crystal?

Current Flow in Semiconductors

In order to understand how transistors operate, there are several new ideas that you must master. First, you must understand how current flows through both N-type and P-type semiconductor materials. Current flow through an N-type material is not too different from current flow through metals, which you have already studied. However, there is quite a difference in the way current flows through a P-type material.

When a P-type material is placed next to an N-type material, we have what is called a junction. The action that occurs at the point of contact between these two different types of materials is extremely important. It is this action that makes the transistor possible.

In this section of the lesson we will study how current flows through N-type and P-type materials. We need to understand current flow through both types of germanium or silicon to be able to understand how a junction works. In a later section we will see how a junction works. Later, we will see what happens in a transistor, which has two junctions.

This section is extremely important, and you should be sure that you understand it completely. Once you understand this material, it will be a simple step to see how transistors can be used to amplify signals.

DIFFUSION

As we have mentioned, adding impurities to pure germanium or silicon adds free electrons or holes. You might at first think that when there is no voltage applied there would be no motion of the free holes and electrons. However, this is not true--as you learned when we discussed intrinsic conduction, there is a certain amount of energy present in the crystal. This energy might be due to the temperature of the crystal, because as we pointed out before, even at room temperature the crystal does have heat energy. Motion of the free holes or electrons due to energy of this type is at random; in other words there is no net movement in any one direction. Holes move one atom at a time, and any hole may move from its starting location to any of the surrounding atoms. This means that a hole may start off in one direction as it moves from one atom to another, and then may move in almost the opposite direction as it moves to still a third atom. Similarly, electron movement is in a random direction; a given electron may move in first one direction and then in another.

When electrons and holes are in motion, the different carriers are moving in different directions. Remember that when there is a hole in one atom, and an electron moves from another atom to fill that hole, a new hole appears in the second atom. In other words, the electron has moved from the second atom to the first, whereas the hole has moved in the opposite direction from the first atom over to the second atom that gave up the electron. The result is that the effective current flow of any one carrier is cancelled by the movement of the other carrier and the resulting current flow in any direction is zero.

This random motion of carriers is called diffusion. It goes on at all times in a crystal whether there is a voltage applied to the crystal or not. Every effort is made in the design of transistors to keep this diffusion as low as possible.

DRIFT

Another type of carrier movement in semiconductors is known as "drift". This is the type of movement that is obtained when a voltage is applied across the crystal. Since the manner in which current flows through N-type and P-type material is different, let's consider them separately.

N-Type Material.

In Fig. 10 we have shown an N-type crystal with a voltage applied to it. The voltage difference supplied by the battery provides a force which makes it easier for the electrons to move in one direction than in the other. In an N-type material, the electrons will be attracted by the positive terminal of the battery. Because in the N-type material the electrons greatly outnumber the holes, they will carry the current.

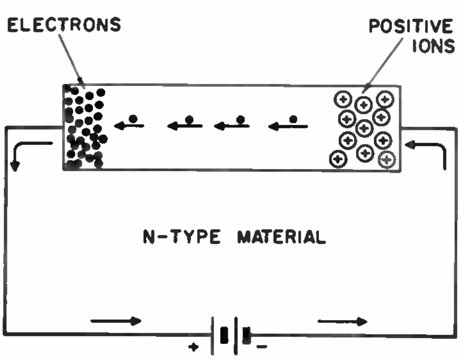


Fig. 10. N-type crystal with voltage applied to it.

When the electrons are attracted by the voltage applied to the positive terminal, they will move towards the positive terminal. When an electron moves away from the covalent bond that produced this free electron, it will leave behind an atom with a positive charge, which we call a positive ion. The electrons moving towards one end of the crystal set up a region that has a local negative charge, as shown in Fig. 10. This negative charge sets up a potential difference between that part of the crystal and the positive terminal of the battery. In other words the attraction of the positive battery terminal causes electrons to bunch up near the end of the crystal connected to the positive terminal. The electrons are drawn from the crystal into the wire connecting the crystal to the positive terminal of the battery by this potential difference.

Meanwhile, the electrons that have left the atoms at the other end of the crystal have left behind positive ions. This sets up a region of positive charge around the end of the crystal connected to the negative terminal of the battery so there will be a potential difference between the negative terminal of the battery and this region of positive charge. This potential difference will pull electrons from the wire into the crystal. These electrons replace the free electrons that were attracted to the positive terminal of the battery.

The number of electrons leaving the crystal at the end connected to the positive terminal of the battery will be exactly equal to the number of electrons entering the crystal at the end connected to the negative terminal of the battery. Since the crystal was electrically neutral be-

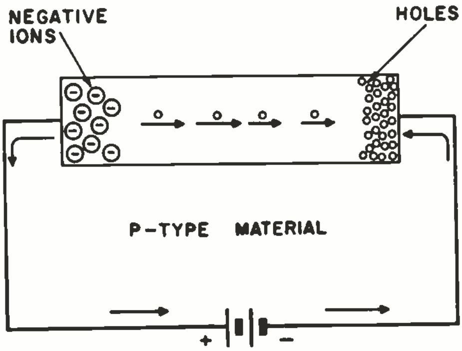


Fig. 11. P-type crystal with voltage applied to it.

fore the battery was connected and the number of electrons in it are constant, the crystal remains electrically neutral.

P-Type Material.

Conduction through P-type material is quite different from conduction through N-type material. In the P-type semiconductor, nearly all of the current is carried by holes. When a battery is applied to a P-type semiconductor, as shown in Fig. 11, the voltage causes the holes to drift towards the negative terminal. They are repelled by the positive potential applied to the one end of the material and attracted by the negative potential applied to the other end. When a hole starts moving away from the end of the material connected to the positive terminal of the battery, it moves because it is filled by an electron attracted from a nearby germanium atom.

When the hole in an accepter type atom is filled with an electron, the atom actually has one electron more than it needs to neutralize the charge on the nucleus. Thus, the atom has a negative charge, or in other words it becomes a negative ion. Negative ions that are formed near the end of the semiconductor that is connected to the positive terminal of the battery build up a region of negative charge at this end of the material. The extra electrons are drawn from these ions by the positive terminal of the battery, and a new hole is formed. These holes then drift towards the end of the semiconductor that is connected to the negative terminal of the battery, and build up a positive charge at this end of the semiconductor. This positive charge attracts free electrons from the external circuit. As a hole is filled with an electron, it disappears.

Thus in the P-type material, we have an electron flow in the external circuit from the negative terminal of the battery to the semiconductor, and from the semiconductor to the positive terminal of the battery. However, in the semiconductor itself, current flow is by means of holes, which drift from the end of the semiconductor that is connected to the positive terminal to the end that is connected to the negative terminal of the battery. Keep this point in mind, that even in the P-type material where conduction within the material is by holes (which are positive carriers) the current flow in the external circuit is by means of electrons and is in the conventional direction from the negative terminal towards the positive terminal of the battery in the external circuit.

There are several important differences between conduction in N-type semiconductors and conduction in P-type semiconductors. In both cases electrons flow from the external circuit into the crystal and then out of the crystal into the external circuit. However, in the N-type crystals, the excess electron produced when a donor atom forms covalent bonds with four germanium atoms is a free electron that can move about in the crystal. However, in the P-type material, the electrons are not free, but can move only to holes. Since a hole can capture an electron from any of its surrounding atoms, it is the hole that is free to move in any direction.

Another important difference between the N-type and the P-type materials is that a free electron moves approximately twice as fast as a hole. This affects the conductivity of the two types of semiconductor material. If we have two crystals, one an N-type and the other a P-type, if the N-type material has the same number of free electrons as the P-type has holes, the N-type will have a lower resistance because the free electrons can move approximately twice as fast as the holes in the P-type material.

SUMMARY

The important thing to remember from this section is that there are two types of carrier movement in semiconductors. The first is called diffusion and is simply a random movement of the carriers in the semiconductor material. The current flow produced by one carrier is cancelled by the movement of the other, and the resultant current flow in any direction is zero. Diffusion is the random motion of electrons or holes in a doped semiconductor due to the energy of the material.

The other type of movement we discussed is called drift. This type of conduction is produced when a potential is connected across a semiconductor. This potential can cause either electrons or holes to move within the semiconductor. In an N-type semiconductor, current flows through the semiconductor because of the movement of the free electrons produced by the donor atoms that have been added to the semiconductor material. In the P-type semiconductor, current flow through the crystal is by means of holes which are produced when an accepter-type impurity is added to the crystal.

In both cases current flow in the external circuit is from the negative terminal of the battery to the crystal and from the crystal to the positive terminal of the battery. In the N-type material, electron flow through the crystal is from the end connected to the negative terminal of the battery to the end connected to the positive terminal of the battery. In the P-type semiconductor, the holes flow from the end of the semiconductor connected to the positive terminal of the battery to the end of the semiconductor connected to the negative terminal of the battery.

The speed with which electrons move through N-type material is about twice the speed with which holes move through P-type material. Thus N-type material has better conductivity than P-type material, which means that N-type germanium will have a lower resistance than P-type germanium.

SELF-TEST QUESTIONS

(t) When an accepter-type impurity is added to a silicon or a germanium crystal, what type of carrier is produced in the crystal?

(u) What is diffusion?

(v) What is the name given to the movement of carriers in a semiconductor material when a voltage is applied across the material?

(w) What are the majority carriers in an N-type material and in what direction do they move when a voltage is applied across the material?

(x) What are the majority carriers in a P-type material and in what direction do they move through the material when a potential is applied across the material?

(y) When current is flowing through a crystal, will the crystal be charged?

(z) Is the rate of travel of electrons through N-type material the same as the rate of travel of holes through P-type material?

(aa) If you had two identical pieces of silicon and one was doped so that it was N-type material and the other doped so that it was P-type material, which would have the lower resistance?

Semiconductor Diodes

Just as there are diode tubes, there are also diode semiconductors. Some diode semiconductors are used as detectors; others are used as rectifiers in power supplies to change ac to pulsating dc. Diodes used as detectors are often referred to as signal diodes. Both germanium and silicon signal diodes are widely used. Diodes used for power rectification are almost exclusively silicon diodes. Relatively small silicon diodes can often handle considerably more current than a large rectifier tube.

A semiconductor diode is made by taking a single crystal and adding a donor impurity to one region and an accepter impurity to the other. This will give us a single crystal with a P section and an N section. Where the two sections meet, we have what is called a junction. Contacts are fastened to the two ends of the crystal so that a simple PN junction diode like the one shown in Fig. 12 is formed. For simplicity in the diagram we have represented the crystal as a box-like structure with one half being P-type material and the other half N-type material with a junction between the two sections.

This type of diode is called a junction diode. The action that takes place at the junction of the P-type crystal and the N-type crystal is what we will be most concerned with now. In order to understand how a junction diode works, you must learn something about the movement of electrons and holes near the junction. The movement of holes and

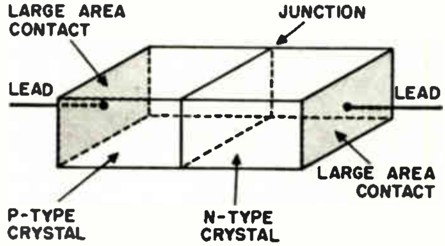


Fig. 12. Simple PN junction.

electrons will form what is called a depletion layer at the junction. Now let us see what the depletion layer is and how it is formed.

DEPLETION LAYER

Remember that in an N-type crystal there are free electrons, and in a P-type crystal there are free holes. Also remember that the electrons and holes are moving about the crystal with a random motion, called diffusion. In the PN junction diode, holes will be moving about in the P section and electrons in the N section. Some of the holes will cross over the junction from the P section into the N section and be filled by a free electron. Similarly, some of the electrons in the N-type material will diffuse across the junction and fill a hole in the P section.

When an atom in the N section loses an electron the atom becomes charged or ionized. It will have a positive charge because it will have one less electron than is needed to completely neutralize the charge on the nucleus. Thus electrons diffusing across the junction to fill a hole on the P side of the junction will leave behind atoms with a positive charge. At the same time, when an electron fills a hole on the P side, the atom will have one more electron than it needs to completely neutralize the charge on its nucleus, and therefore that atom will have a negative charge. Similarly, holes diffusing from the P side of the junction over into the N side will leave behind atoms with a negative charge. When the hole moves over to the N side, it will mean the atom into which it moves will have an electron missing and therefore it will assume a positive charge. When the hole leaves the P side of the junction because it has been filled by an electron, the atom that gains the extra electron will have a negative charge.

As a result of this diffusion across the junction, a region will build up around the junction called the depletion area. On the P side of the junction there will be an area where the holes are missing. On the N side of the junction there will be an area where electrons are missing; thus we get the name depletion layer.

The missing holes on the P side of the junction will result in a negative charge on the P side and the missing electrons on the N side will produce a positive charge on the N side of the junction. The negative charge on the P side of the junction will build up until it has sufficient amplitude to prevent any further electrons from the N side from crossing the junction to the P side. Remember that the negative charge built up on the P side of the junction will repel electrons from the negative side. Similarly, the positive charge built up on the N side of the junction will prevent holes from the P side from crossing the junction into the N-type material. Thus this area, which is called the depletion layer because it is short holes on one side and electrons on the other side, is also sometimes called the barrier layer, because the charges built up form barriers to prevent any further diffusion of holes or electrons across the junction. It is also sometimes called a potential barrier because a negative potential is built up on the P side of the junction and a positive potential is built up on the N side of the junction.

The action taking place at the junction is quite important and is illus-

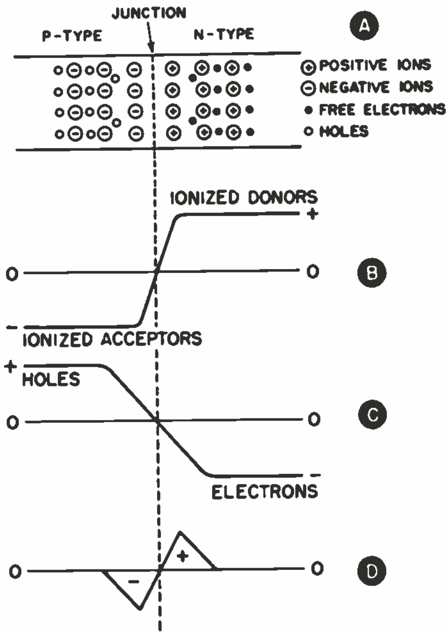


Fig. 13. (A) locations of ions and carriers at a PN junction; (B) charges at junction due to ionized impurity atoms; (C) carrier charges available; (D) resultant charges.

trated in Fig. 13A. On the P side of the junction we have shown ionized atoms that have a negative charge because the holes in these atoms have been filled by electrons. The holes have escaped and travelled or diffused across the junction into the N-type material. On the N side of the junction we have shown atoms that are ionized and have a positive charge. These atoms have a positive charge because they have lost electrons. These electrons have diffused across the junction into the P-type material. Thus we have a charged area at the junction. The negative charge on the P side of the junction prevents any further movement of electrons from the N-type material across the junction into the P-type material, and the positive charge on the N side of the junction prevents any further movement of holes from the P-type material across the junction into the N-type material.

The charge on the ions is shown in Fig. 13B. Notice that on the P side of the junction the atoms that have lost holes by gaining electrons have a negative charge. At the junction the potential drops to zero and then reverses on the N side where the ionized atoms have a positive charge because they have lost electrons.

In Fig. 13C we see the carrier charges which are available to neutralize the ionized atoms. At some distance from the junction there are holes with a positive charge. However, as we approach the junction, the concentration of these holes decreases because they are repelled away from the junction by the positive ions on the N side of the junction. On the N side of the junction at some distance from the junction we have many electrons available, but as we approach the junction, the charge drops to zero because these electrons are repelled away from the junction by the negative ions on the P side of the junction.

The resultant charges on the crystal are shown in Fig. 13D. As before, the crystal will have a tendency to remain neutral, or in other words not to have any charge. Some distance from the junction the atoms will have exactly the correct number of holes and electrons so that the net charge on the atoms is zero. As we approach the junction, the negative ions on the P side will result in an area in the crystal that has a negative charge. As we move closer to the junction, the charge will drop to zero so that at the junction itself the net charge on the atoms is zero. Then the charge builds up in a positive direction on the N side of the junction due to the ionized atoms that have lost electrons. As we move away from the junction we again reach a region where the atoms have exactly the correct number of electrons to neutralize the charges on the nucleus so the net charge in that area will be zero.

So far we have been discussing only the action of the majority carriers at the PN junction. However, there is one other important point we must consider in order to completely understand what happens at the junction. You will remember some time ago that we mentioned that holes and electrons are in a continuous state of motion in the crystal due to the energy of the crystal. For example, even at room temperature, the crystal contains a certain amount of heat energy and this energy is sufficient to cause motion of both electrons and holes. In the N-type material an electron will leave an atom creating a hole. This hole will be filled by an electron from another atom. Thus we have the continual formation of hole-electron pairs. Away from the junction, this formation of hole-electron pairs does not have any effect on the carrier concentration in the crystal. In other words, the holes will remain the majority carriers in the P-type region, and the electrons will remain the majority carriers in the N-type side of the crystal.

However, as we mentioned previously, both holes and electrons are involved in conduction at all times. There are minority carriers in both regions - holes in the N region and electrons in the P region. The holes produced in the N region near the junction will be attracted by the negative ions on the P side of the depletion layer at the junction and pass across the junction. These holes will tend to neutralize the ions on the P side of the junction. Similarly, free electrons produced on the P side of the junction will pass across the junction, and neutralize positive ions on the N side of the junction. This is an example of intrinsic conduction, conduction due to the formation of hole-electron pairs, and as we mentioned, this type of conduction is undesirable.

Now let us consider what happens due to the minority carriers crossing the junction. Holes crossing the junction from the N-type material to the P-type material tend to neutralize the positive ions on the P side of the junction. Similarly, electrons traveling from the P side of the junction to the N side of the junction tend to neutralize the negative ions on the N side of the junction. This flow of minority carriers across the junction weakens the potential barrier in the region around the atoms they neutralize. When this happens, the majority carriers are able to cross the junction at the location of the neutral atom. This means that the holes from the P side will cross over to the N side, and electrons from the N side will cross over to the P side.

The result is that we have both holes and electrons crossing the junction in both directions. The hole that crosses from the N side to the P side due to intrinsic conduction permits a hole to cross from the P side to the N side by diffusion. Similarly, an electron that crosses the junction from the P side to the N side due to intrinsic conduction permits another electron to go from the

N side to the P side by diffusion. The result of the holes and electrons crossing the junction in both directions is that these movements cancel each other, and the charge on the atoms at the junction remains the same. This movement of holes and electrons in both directions contributes nothing towards the net charge or current flow through the junction. However, the flow across the junction will produce a certain amount of heating; it will in effect use up a percentage of the total capacity of the junction to pass current so that the net result is to reduce the amount of useful current the diode can pass.

BIASED JUNCTIONS

If a battery is connected to the ends of a PN junction diode, the battery potential will bias the junction. If we connect the battery so that its polarity aids the flow of current across the junction, we call it a "forward-biased junction", whereas if we connect the battery so that the polarity opposes the flow of current across the junction, we say that it is a "reverse-biased junction". In both cases there will be some current flow through the junction, but as you might expect, with forward bias the current flow will be higher.

In order to understand how transistors work, you must understand both conditions of bias. You will study each condition separately, because the action that occurs at the junction is quite different in the two cases. In the operation of transistors both types of bias are used, and therefore it is important that you understand what happens in each case.

Forward Bias.

When we connect a battery to a junction diode with the polarity such that it aids the movement of majority carriers across the junction, we say that the diode is forward biased. A forward-biased junction is shown in Fig. 14. Here the positive terminal of the battery is connected to the P-type section and the negative terminal of the battery is connected to the N-type section. Now let us consider what happens to the depletion layer at the junction of the P and N-type material when the battery voltage is applied.

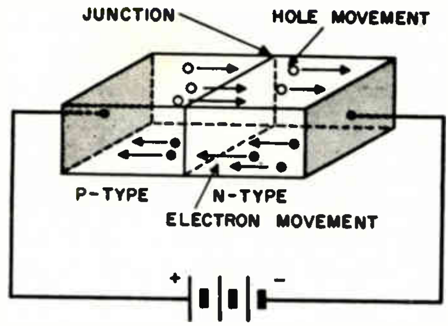


Fig. 14. Forward-biased junction.

The positive voltage connected to the end of the P-type crystal will repel holes towards the junction and attract electrons from the negative ions near it. The combination of holes moving towards the junction to neutralize charged negative ions on the P side of the junction and electrons being taken from the negatively charged ionized accepter atoms tends to neutralize the negative charge on the P side of the junction.

On the N side of the crystal, the negative terminal of the battery repels electrons towards the junction. These electrons tend to neutralize the positive charge on the donor atoms at the N side of the junction. At the same time the negative potential at the N side of the crystal attracts holes away from the charged positive ions on the N side of the junction. Both of these actions tend to neutralize the positive charge on the donor atoms at the junction.

The effect of the battery voltage is to reduce the potential barrier at the junction and allow more majority carriers to cross the junction. This means that we will have more electrons flowing from the N-type material across the junction to the P-type material and to the positive terminal of the battery and more holes traveling from the P-type material across the junction to the N-type material and towards the end of the crystal connected to the negative terminal of the battery. You know that we already had a certain number of intrinsic minority carriers crossing the junction, but now the majority carriers outnumber them, so there will be a steady current flow from the negative battery terminal, through the N-section, across the junction and through the P-section, to the positive battery terminal.

Placing a forward bias on a junction diode drives majority carriers back into the depletion layer and allows conduction across the junction. If the battery voltage is increased, more carriers will arrive at the junction and the current flow will increase. Eventually, if we continue to increase the battery voltage, we will reach a point where all the charges at the junction are neutralized. When this happens, the holes will fill the P-type region right up to the junction; electrons will fill the N-type region up to the junction; and the only limit to current flow through the diode will be the resistance of the material on the two sides of the junction.

It is important for you to remember that in a forward biased junction conduction through the crystal will be by the majority carriers. Any intrinsic conduction across the junction will be by minority carriers and this will subtract from the total current flow across the junction. Increasing the forward bias will increase the current flow across the junction until the point is reached where all the charges at the junction are neutralized, at which time the potential barrier will disappear, and current flow across the junction will be unhindered by any potential across the junction.

Reverse Bias.

If we reverse the battery connections we will have what is known as reverse bias. This condition is shown in Fig. 15.

With a reverse bias applied to a junction diode, the negative terminal of the battery will be connected to the P-type section, and will attract holes away from the junction, and increase the shortage of holes on the P side of the junction. At the same time the positive terminal of the battery is connected to the N-type section of the crystal and this

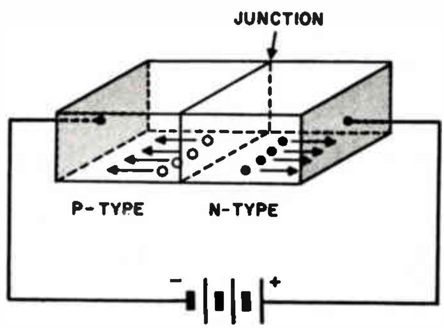


Fig. 15. Reverse-biased junction.

terminal will attract electrons away from the junction and increase the shortage of electrons on the N side of the junction. This movement of holes and electrons away from the junction will in effect result in an increased potential barrier at the junction. The increase in potential barrier occurs because there will be fewer holes on the P side of the junction to neutralize the negative ions and fewer electrons on the N side to neutralize the positive ions formed on this side of the junction. The increase in potential barrier will help prevent any further current flow across the junction due to majority carriers.

The current flow across the barrier, however, is not zero because we will still have minority carriers crossing the junction. Holes forming in the N side of the depletion layer will be attracted by the negative potential applied to the end of the P-type section of the crystal, and electrons breaking loose from their nuclei in the P side of the depletion layer will be attracted by the positive voltage applied to the end of the N-type section of the crystal.

We had this situation when there was no bias applied to the junction. Holes from the N side would cross over to the P section, and electrons from the P side would cross over to the N section. However, when there was no bias applied to the crystal, these minority carriers would neutralize ions near the junction and allow the majority carriers to cross the junction. However, since the minority carriers are now attracted away from the junction by the potential applied to the crystal, all of the minority carriers do not remain near the junction to neutralize charged atoms so they no longer allow the passage of an equal number of majority carriers in the opposite direction. This means that the flow of minority carriers across the junction is not fully offset by a flow of majority carriers in the opposite direction. Therefore, there will be a small current flow across the junction due to the minority carriers crossing the junction. This current flow is very small and nearly constant at all normal operating voltages in signal diodes and power rectifier diodes. However, as you will see later, there are certain types of diodes where this reverse current can increase quite rapidly even at low voltages.

It is important to realize that when a reverse bias is applied to a junction diode, the bias increases the potential difference across the junction and makes it more difficult for majority carriers to cross the junction. However, some minority carriers will still cross the junction with the result that there will be a small current flow across the junction due to the minority carriers.

COMPARISON OF JUNCTION DIODES AND VACUUM TUBES

Although the operation of junction diodes designed for use as rectifiers is quite different from the operation of vacuum tubes, they can perform identical tasks and therefore some comparison of the most important characteristics of both is in order.

When there is no voltage applied to a junction diode, the net current flow across the junction is zero. On the other hand, in a vacuum tube, even though there may be no voltage applied to the plate of the tube, some of the electrons will leave the cathode with sufficient velocity to travel across the space between the cathode and the plate and strike the plate. This will result in a small current flow from the cathode to the plate of the tube even though there may be no voltage applied to the plate of the tube.

We can consider applying a positive voltage to the plate of a vacuum tube and a negative voltage to the cathode as being similar to placing a forward bias on a junction diode. Under both circumstances there will be a current flow through the diode. In this respect the two are similar.

When the voltages applied to the diode vacuum tube are reversed so that there is a negative voltage applied to the plate and a positive voltage to the cathode, there will be no current flow at all through the tube. The negative potential on the plate of the tube will repel electrons from the plate. This reverse voltage situation is similar to a reverse bias across a junction diode. However, when we place a reverse bias across a junction diode, there will be some current flow across the junction due to the conduction by minority carriers. As long as the breakdown voltage of the junction diode is not exceeded, this current will be very small and almost constant. In a good diode, it is so small it can be ignored.

We can summarize the characteristics of diode vacuum tubes and junction diodes as follows: with forward bias both the tube and junction diodes will conduct. With reverse bias, the tube will not pass current; the junction diode will pass a small current. With no bias, the tube will pass a small current; the junction diode will not.

ZENER DIODES

In junction diodes designed for use as rectifiers we must be careful not to exceed the rated reverse voltage of the diode. In other words, if we place too high a reverse bias across the junction, the junction will break down, a very high current will flow across the junction for a short while, and the diode will be destroyed. However, in some diodes we make use of this reverse current due to minority carriers. In diodes of this type both the P section and the N section are doped quite heavily. The junction between the P section and the N section is considerably larger than the junction of the rectifier-type diode, so that when the diode begins to pass current in the reverse direction, it can pass it over a larger area and thus avoid destroying the diode. This type of diode is used as a voltage reference and is referred to as a voltage-reference diode or a Zener diode.

In the Zener diode, the current remains small with low reverse voltages. At a certain voltage, called the breakdown voltage, the current will increase rapidly with any further increase in voltage.

The breakdown voltage can be varied by varying the diode material and construction. Zener diodes can be made with a breakdown voltage as low as 1 volt, up to breakdown voltages of several hundred volts. The current that can pass through any Zener diode before the diode will be damaged will depend upon the junction area and the methods used to keep the diode cool.

In a circuit where a Zener diode is used, it will be used as a voltage reference or a voltage regulator. As the reverse voltage across the diode

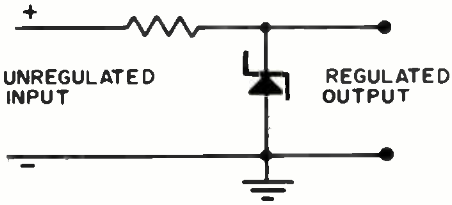


Fig. 16. Circuit using Zener diode as a voltage regulator.

increases, a very small reverse current will flow, and the value of this current will remain essentially constant until the breakdown voltage is reached. At that voltage, any further increase in voltage will result in a large increase in current across the junction. This large increase in current tends to produce a voltage drop in other components in the circuit so that the voltage across the diode will remain essentially constant. Thus a Zener diode can be used as a voltage regulator in a circuit such as shown in Fig. 16. If the input voltage tends to rise, the current through the diode will increase. This increase in current will increase the voltage drop across the resistor so that the output voltage will remain essentially constant. The fact that with even a very small increase in voltage the current through the diode will increase substantially means that the voltage across the diode will remain almost constant.

In some other applications the diode may be used as a reference voltage. This simply means that a Zener diode with a given breakdown voltage is used in a circuit like Fig. 16. The voltage applied to the diode will remain essentially constant. In some circuits, we may compare the voltage across a resistor or another part with the voltage across the Zener diode.

TUNNEL DIODES

A tunnel diode is a highly doped junction diode made either of germanium or gallium arsenite. Both the N region and the P region of the diode are very highly doped. As a result of the high doping, the depletion region around the junction is extremely narrow. Because of the narrow depletion region, holes and electrons can cross the junction by more or less tunneling from one atom to another. The exact action of the charges crossing the junction is somewhat difficult to visualize, but the characteristics of the tunnel diode are comparatively simple. With a reverse bias across the junction, current across the junction increases quite rapidly. If the reverse bias is dropped to zero the current will drop back to zero. If the forward bias, starting at zero and gradually increasing, is placed across the junction, current flow across the junction will increase at essentially the same rate as it increased with a negative bias. The current across the junction will increase quite rapidly as the forward bias is increased until a rather sharp peak is reached. If the forward bias is increased beyond this point, then the current begins to decrease.

As the forward bias is increased still further, the current across the junction decreases, forming a curve such as shown in Fig. 17. This decreasing current, with increasing voltage, results in a negative resistance characteristic. It might be difficult to visualize what a negative resistance is, but you will remember from Ohm's Law that re-

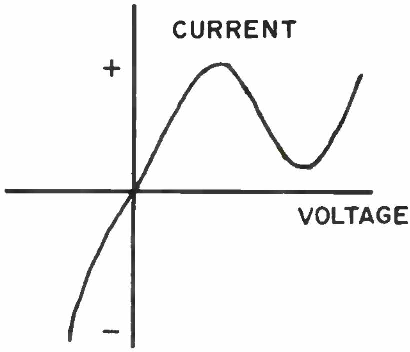


Fig. 17. Voltage-current relation in a tunnel diode.

sistance is equal to voltage divided by current. In a circuit where we have resistance, if the resistance is constant and the voltage increases, then the current must increase; and similarly if the voltage decreases, the current must decrease. Here in the tunnel diode we have a region where the opposite happens. If the voltage increases, the current decreases and if the voltage decreases, the current increases. Thus we have something in the circuit that is giving us the opposite effect of resistance; we call this negative resistance. You will remember that resistance in a circuit introduces losses. It is the resistance in a resonant circuit that prevents a resonant circuit from continuing to oscillate once it has been excited into oscillation. However, if we can put something with negative resistance in the circuit, for example a tunnel diode, since it has the opposite effect of resistance, then the circuit should continue to oscillate. Tunnel diodes can be used for this purpose.

At the present time, tunnel diodes have not appeared in the commercial entertainment-type equipment. However, it is probably just a matter of time until they are used; therefore you should at least have some basic knowledge of what the tunnel diode is.

P-I-N DIODES

The p-i-n diode, which is an abbreviation for positive-intrinsic-negative, is a new diode which is used in a somewhat different manner from the diodes you have studied previously. Rather than being used as a detector or rectifier, this diode is used primarily as a variable resistor. It is a special type of diode, and its resistance can be controlled by applying a dc bias to it. With a reverse bias across the diode, it has a very high resistance. With no bias, its resistance drops to about 7000 ohms, and with forward bias it drops to a comparatively low value.

The diode is particularly useful in circuits where the strength of a signal must be controlled. Its first commercial use has been in fm equipment and it is used in order to prevent extremely strong fm signals from causing overloading in the fm receiver. A dc bias is applied to the diode and the amplitude of the bias depends upon the strength of the signal. When a very strong signal is applied, the reverse bias applied to the diode increases so that the resistance of the diode increases. This reduces the strength of the signal fed to the mixer and i-f stages in the receiver and thus prevents overloading, particularly in the last stage of the receiver.

At this time p-i-n diodes are not widely used in commercial applications, but you should be aware of how the diode is used, because it is quite likely that it will be widely used in the future.

POINT-CONTACT DIODE

Another semiconductor diode is the diode detector used in many TV receivers. This detector is a point-contact diode. A cut-away view of a point-contact diode is shown in Fig. 18 along with the schematic symbol.

The point-contact diode is made of a small piece of either N-type or P-type germanium or silicon. N-type germanium or silicon is used more often than P-type. In manufacturing a diode of N-type material, a large contact is fastened to one side of the crystal. A thin wire, called a catswhisker, is attached to the other side of the crystal. When the catswhisker is attached to the N-type crystal, a small region of P-type material is formed around the contact as shown in Fig. 19. Thus we have a PN junction that performs in much the same way as the junctions we have already discussed.

The characteristics of the point-contact diode under forward and reverse bias are somewhat different from those of the junction diode. With forward bias the resistance of the diode is somewhat higher than that of a junction diode. With reverse bias the current flow through a point-contact diode is not as independent of the voltage applied to the crystal as it was in the junction diode. In spite of these disadvantages, the point-contact diode makes a better detector than the junction diode, particularly at high frequencies because the point-contact diode has a lower capacity than the junction diode.

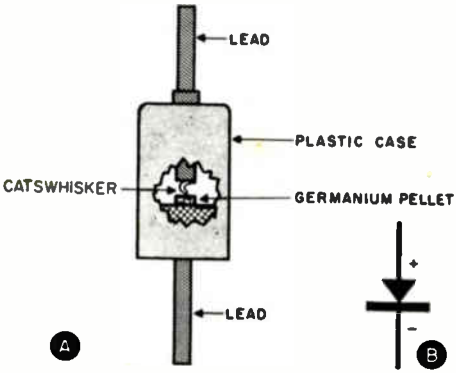


Fig. 18. (A) cut-away view of crystal rectifier. (B) schematic symbol of it.

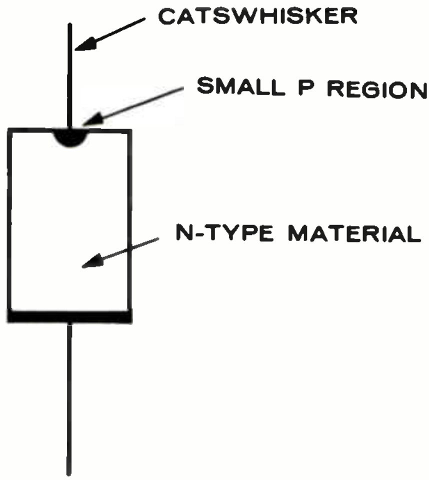


Fig. 19. Sketch of a point-contact diode showing where the P-type germanium is formed around the catswhisker.

SUMMARY

Again this is a very important section, so you should review it before going on to the next section. Make sure you understand what the depletion layer is and why it is formed. Also be sure you understand the movement of holes and electrons across a PN junction when no voltage is applied to the junction.

Current flow across the junction with both forward and reverse bias is important. With forward bias current flow across the junction is by majority carrier, and with reverse bias it is by minority carrier.

SELF-TEST QUESTIONS

(ab) What are the two principal uses of semiconductor diodes?
(ac) What is the depletion layer?
(ad) What do we mean by the potential barrier?
(ae) Does the crystal develop an overall charge as a result of diffusion across the junction?
(af) Do the minority carriers crossing the junction have any adverse effect on the diode?
(ag) What do we mean when we say the junction is forward biased?
(ah) What do we mean when we say a junction is reverse biased?
(ai) What is the difference between vacuum-tube diodes and semiconductor diodes insofar as current flow through the diode is concerned when no voltage is applied?
(aj) What is the difference between current flow in a semiconductor diode and a vacuum tube under reverse voltage conditions?
(ak) What is a Zener diode, and what is it used for?
(al) What is a tunnel diode?
(am) What is a p-i-n diode?

Semiconductor Triodes

Even though a junction diode will pass current in both directions, it passes current in one direction much better than it does in the other, and therefore it can be used as a detector. The tunnel diode can be used as an oscillator, and in some special circuits as an amplifier; however, its usefulness in these applications is limited. In most cases, the semiconductor diode is like the vacuum-tube diode; it is more or less useless insofar as amplifying a signal is concerned. In order to amplify a signal, a three element semiconductor is needed. Three element semiconductors that are capable of amplification are called transistors.

There are a number of different types of transistors in use today. The characteristics of the different types vary appreciably, but if you understand the operation of one type, you can understand how the others work without too much difficulty. We started our explanation of. semiconductor devices with a junction diode, so we will start our study of triode semiconductors with a study of the junction transistor. You'll find that most of the transistors you will study operate in a manner similar to the basic junction transistor. The most notable exception to this is the field-effect transistor which you will study later in this lesson.

JUNCTION TRANSISTORS

Both germanium and silicon are used in the manufacture of junction transistors. A triode-junction transistor is made up of single semiconductor crystals with three different regions. The center region is made up of one type of germanium or silicon, and the two end regions are made up of the other type of germanium or silicon. In other words, in one type of junction transistor the center has had accepter-type region impurities added and the two end regions have had donor-type impurities added. In the other type of junction transistor, the center region has had donor-type of impurities added and the two end regions have had accepter-type impurities added.

The center region of the transistor is called the base. This is usually a comparatively thin region. One of the end sections is called the emitter and the other end section is called the collector.

If the center section of the crystal has been treated with donor-type impurities, then the center section becomes N-type germanium in the case of a germanium transistor or N-type silicon in the case of the silicon transistor. In this case, the two end sections will be treated with accepter-type impurities and they will both become P-type germanium or P-type silicon. We call this type of transistor a PNP transistor. We can have both germanium PNP and silicon PNP transistors. An example of this type of junction transistor is shown in Fig. 20A along with the schematic symbol used to identify it.

The other type of junction transistor is an NPN type. This type of transistor is produced by treating the center section with an accepter-

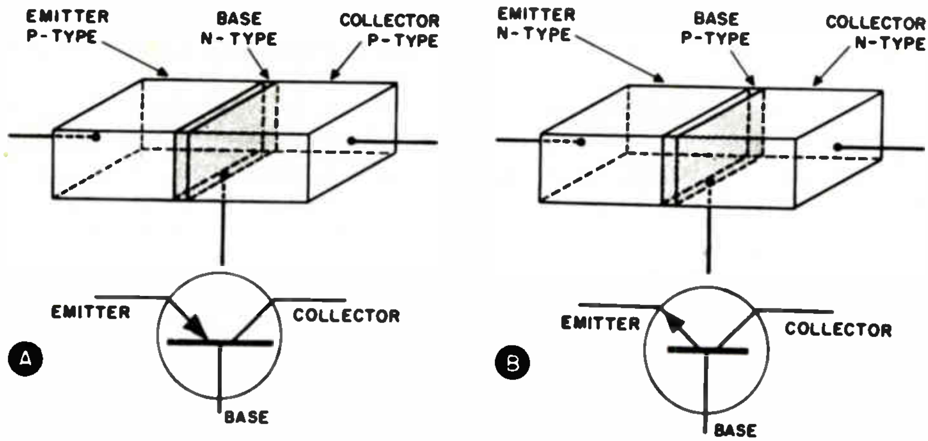


Fig. 20. (A) shows a PNP junction transistor and its schematic symbol. (B) shows an NPN junction transistor and its schematic symbol.

type impurity to produce P-type germanium or silicon, and the two end sections with a donor-type of impurity to produce N-type germanium or silicon. This is the type of junction transistor shown in Fig. 20B along with the schematic symbol for it. As in the case of the PNP transistor, we can have either an NPN germanium transistor or an NPN silicon transistor.

Notice that the schematic symbols for the PNP transistor and the NPN transistor are different. In the PNP transistor the arrow used to represent the emitter points down toward the base, whereas in the NPN transistor the arrow on the emitter points up away from the base. Thus on a schematic diagram of a piece of equipment using junction transistors, you can tell from the direction in which the arrow is pointing whether the transistor is a PNP or an NPN transistor.

There are several different ways of manufacturing junction transistors, and often you hear these transistors referred to by the manufacturing method used. For example a junction transistor might be called a grown junction, a fused junction, an alloy junction or a diffused junction. All these names simply describe the manufacturing process used to make the transistor. They are all junction transistors and all operate on the same basic principle. This does not mean to imply that the characteristics of all junction transistors are the same; they are not. There are wide differences in characteristics just as there are in various triode vacuum tubes.

To understand how junction transistors work, you must understand junction diodes. Because it is so important that you understand the formation of the depletion layers at the junctions we will review the explanation of what happens at the junctions in explaining the junction transistors. The big difference between the junction transistor and the junction diode is that in the transistor there are two junctions close together in the same crystal. One of these junctions is biased in the forward direction and the other in the reverse direction and the presence of one junction affects the operation of the other.

PNP TRANSISTORS

In transistor operation, the emitter-base junction is always biased in the forward direction and the collector-base junction is biased in the reverse direction. Each of the two junctions by itself behaves just like the PN junction already described.

Let us consider what happens in the PNP transistor before any voltages are applied to the transistor.

At the junctions, holes from the P-type emitter section and the P-type collector diffuse across the junctions into the base. At the same time, electrons from the base diffuse across the junctions into both the emitter and the collector. The holes diffusing into the base place a positive charge on the atoms near the junctions. Similarly the electrons diffusing from the base into the emitter on one side of the base and the collector on the other side of the base place a negative charge on the atoms on the emitter and collector sides of the junctions. These charged atoms, which are called ions, will repel electrons and holes from the region of the junctions. The positively charged ions in the base will repel holes in the P sections away from the junctions. Similarly, the negatively charged ions in the emitter and collector will repel electrons away from the junctions in the base. Thus we have two depletion layers formed, one at the emitter-base junction and the other at the base-collector junction.

You will remember that when we discussed the junction diode, we mentioned that hole-electron pairs will be formed in the depletion region. The minority carriers formed in each section can cross over the junction. For example, the electrons released in both the emitter and the collector regions will cross the junctions into the base. These electrons will neutralize a few ions in the base region. When these ions are neutralized they will allow majority carriers from the emitter and collector to cross the junctions. In other words, there will be holes from the emitter and holes from the collector crossing the junctions into the base. Similarly, holes, which are the minority carriers in the base region (and are formed in the depletion layer) will cross the junctions into the emitter and collector. When these holes cross the junctions they will neutralize some of the nega-

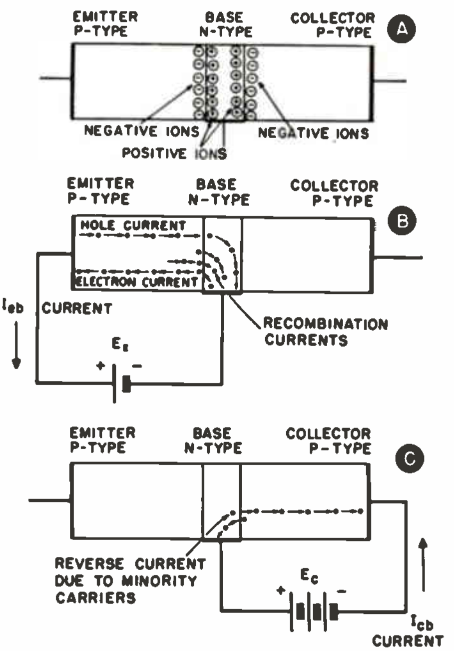


Fig. 21. (A) the formation of ions at the junctions of a PNP transistor. (B) the current flow in the emitter-base circuit and (C) in the collector-base.

tively charged ions in the emitter or collector region and allow some electrons to flow from the base into either the emitter or the collector. Thus, because of the intrinsic conduction due to hole-electron pairs being formed in the depletion region there will be some flow of carriers across the junction. However, the flow of majority carriers across the junction will be exactly equal to the flow of minority carriers across the junction so that the net current flow across each junction will be zero.

The potential barriers formed at the junction regions are shown in Fig. 21A. Notice that the charges formed at the junction are similar to those formed at a junction diode; we simply have two junctions to consider in a transistor.

Now when we place a forward bias between the emitter and the base we have an arrangement like that shown in Fig. 21B. Here the positive voltage applied to the end of the P-type emitter repels holes towards the junction. These holes tend to neutralize the negative charge on the ions on the emitter side of the junction. The holes are formed at the end of the P-type section by electrons being taken out of this section by the positive potential applied to it. At the same time the positive potential applied to the emitter attracts the electrons that have given the ions on the P side of the junction their negative charge. This also weakens the negative charge on the emitter side of the junction.

At the base, which is connected to the negative side of the battery, the holes will be attracted toward the negative terminal of the battery, and electrons will be pushed towards the depletion layer. The pulling of holes away from the depletion area and pushing electrons into the depletion area tends to neutralize the charge on the base side of the junction.

The net effect of biasing in a forward direction is to neutralize the charges on each side of the junction and allow current to flow across the junction. Current flow is by majority carriers: electrons from the N-type base region and holes from the P-type emitter region.

Thus in the emitter-base circuit we have electrons flowing from the negative terminal of the battery to the base, through the base, across the junction, and through the emitter to the positive terminal of the battery. At the same time we have holes being produced because electrons are being pulled out of the P-type emitter by the positive potential applied to it. The holes will move through the emitter, across the junction into the base and to the point where the base is connected to the negative terminal of the battery. At this point they will pick up electrons and disappear.

Not all the electrons going from the base to the emitter will reach the positive terminal of the battery. Some of these electrons will recombine with holes in the emitter. Similarly, some of the holes traveling from the emitter into the base will pick up an electron in the base. This current flow across the junction is called a recombination current, and the transistor is designed to keep this current as low as possible. In other words we want the holes and electrons crossing the junction to reach the terminals connected to the battery.

Now let us consider the other junction, the base-collector junction. This junction is reverse biased as shown in Fig. 21C. Here again we have a depletion layer at the junction. Also we have minority carriers being formed in the depletion layer. However, holes that are formed in the base will cross the junction and then instead of neutralizing a negatively charged atom near the junction in the collector, these holes will be attracted by the negative potential applied to the collector. Similarly, electrons formed in the depletion layer of the P-type collector will cross the junction and be attracted by the positive potential applied to the base. Thus we have a current flow due to the minority carriers. Electrons in the depletion layer of the collector section will cross the junction and flow through the base to the positive terminal of the battery. Meanwhile electrons from the negative terminal of the battery will fill holes that are moving from the base, across the junction, and through the collector to the negative terminal.

Thus you can see that while we have a current by majority carriers due to the forward bias applied between the emitter and the base, we also have a small current flowing through the base-collector circuit by minority carriers due to the reverse bias applied between the base and the collector. Now let us see how the two junctions affect each other.

A transistor with both biased junctions is shown in Fig. 22. Here we have a number of different currents flowing. In the emitter-base circuit we have current flowing due to the forward bias applied between these two. Electrons will flow from the negative terminal of the battery into the base, across the junction, and through the emitter to the positive terminal of the battery. We will also have some holes formed in the P-type emitter section due to electrons being pulled out of this section by the positive terminal of the battery. Some of these holes will cross the junction into the N-type base where they will pick up an electron and disappear. This current is called the recombination current.

Many of these holes will cross the base and flow through the collector, because the negative terminal of the battery connected between the base and collector will attract them. This movement of holes accounts for most of the current flow in the emitter and collector circuits. Remember that holes are being continually formed in the P-type emitter because electrons are

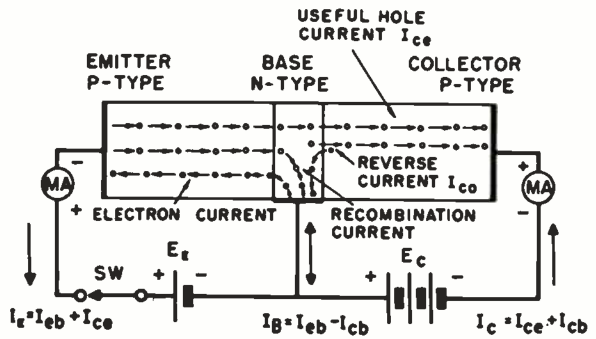


Fig. 22. Current flow and carrier movement in a PNP junction transistor.

being pulled out of the emitter by the positive potential applied to it. These holes will continually move through the emitter, and into the base. Here some of them will combine with electrons and disappear, but the majority of them will flow through the collector to the negative terminal of the collector where they will be filled by electrons and disappear.

Another current that will flow is reverse current I_{CO} that flows in the base-collector circuit. This is due to the formation of minority carriers in the depletion layer.

Thus we have four currents flowing in the PNP junction transistor. The largest of these currents is due to the movement of holes from the emitter through the base into the collector to the negative terminal of the battery connected to the collector. We have in addition to this current three small currents flowing. We have the current due to the electron movement from the negative terminal of the emitter-base battery into the base, across the junction and through the emitter to the positive terminal of this battery. We have the recombination current due to holes combining with electrons in the base, and we have the reverse current due to hole-electron pairs being formed in the depletion layer of the base-collector junction. The directions of the different movements of holes and electrons are marked in Fig. 22.

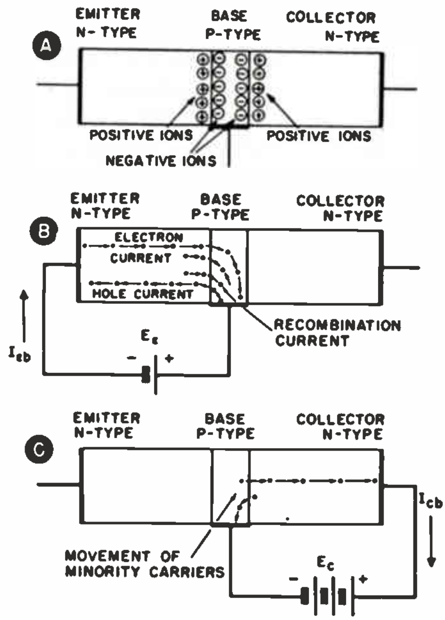


Fig. 23. (A) the formation of ions at the junctions of an NPN transistor. (B) the current flow in the emitter-base circuit and (C) in the collector-base circuit.

NPN TRANSISTORS

Although the operation of the NPN transistor is somewhat different from that of the PNP type, if you understand how the PNP transistor works, you should have no difficulty understanding the NPN. Again, we can start our study of this type of transistor by considering what happens at the junctions, remembering that the action at the junction is similar to the action we studied at the simple PN diode junction.

Let us first consider the action of the holes and electrons before any voltages are applied to the transistor. The charges that will be built up are shown in Fig. 23A. Remember that holes from the base will diffuse across both junctions into the emitter and the collector. Similarly electrons from the emitter and electrons from the collector will diffuse across the junctions into the base. The holes and electrons diffusing across the junctions will charge atoms near the junction. Holes crossing the junctions into the emitter and the collector will ionize the atoms on the emitter and collector sides of the junctions so that they will have positive charges. Similarly, electrons diffusing across the junctions into the base will ionize atoms in the base near the junctions so that they will have negative charges. Thus we will have potential barriers at the junctions. This is the same kind of potential barrier that we found existed across the PN junction in a diode.

The positively charged ions on the emitter and collector sides of the junctions will force holes in the base away from the junction. Similarly the negatively charged atoms on the base side of the junctions will force electrons in the emitter and collector away from the junction so that at the junctions we have a depletion layer.

Now let us consider what happens when we apply a forward bias between the emitter and the base by connecting a battery between the two as shown in Fig. 23B. Notice that the negative terminal of the battery is connected to the end of the emitter, and the positive terminal is connected to the base.

Now, several things happen. The negative potential applied to the emitter will force electrons toward the junction. At the same time the negative potential will attract holes away from the junction. Both of these actions tend to neutralize the positively charged ions on the emitter side of the junction. At the same time the positive terminal of the battery that is connected to the base will attract electrons away from the negatively charged atoms on the base side of the junction. In addition, the positive potential will repel holes towards the junction so that these two actions tend to neutralize the charge on the base side of the junction.

Once the potential barrier at the junction is weakened, electrons can flow from the negative side of the battery into the emitter, through the emitter, and across the junction into the base and from the base to the positive side of the battery. At the same time the positive terminal of the battery can extract electrons from the base, forming holes. Holes are then repelled toward the junction, across the junction, and through the emitter toward the end of the emitter that is connected to the negative terminal of the battery. Here the holes will pick up electrons and disappear. Thus we have a current flow through the emitter-base circuit as shown in Fig. 23B.

Now let us consider what happens when we apply a reverse bias between the base and the collector. Here the negative potential applied to the base will pull holes away from the junction, and the positive potential applied to the collector will pull electrons away from the junction. Thus the negative charge on the base side of the junction will be increased, and the positive charge on the collector side of the junction will be increased so that the potential barrier at the junction will be increased. This will prevent any current flow through the base-collector circuit due to the majority carriers.

At the same time electrons, which are minority carriers, will break loose from their nuclei in the depletion layer on the base side of the junction and will be attracted by the positive potential applied to the collector. They will cross the junction and flow through the collector to the terminal connected to the positive side of the battery, as shown in Fig. 23C. At the same time holes formed on the collector side of the junction in the layer will be attracted by the negative terminal of the battery, and hence will cross the junction and flow over into the base and toward the negative terminal of the battery. Here they will pick up an electron and disappear.

Thus we will have a current flow in the base-collector circuit due to the minority carriers. This is the same situation that we had in the reverse biased base-collector circuit of the PNP transistor.

Now let us see what happens when bias voltages are applied across both junctions of the complete NPN junction transistor as shown in Fig. 24. Considering first the emitter-base circuit, we have electrons flowing from the negative terminal of the battery to the N-type emitter. Here the electrons flow through the emitter, across the junction, and into the base. Some of these electrons reaching the base will recombine with holes in the base. This is called the recombination current. However, the majority of the electrons reaching the base will be attracted by the positive potential applied to the collector and hence will flow through the base across the base-collector junction and through the N-type collector to the positive terminal of the battery in the base-collector circuit.

At the same time the positive terminal of the battery in the emitter-base circuit is connected to the base, and this potential will pull electrons out of the P-type base, producing holes. These holes will then cross the junction into the emitter, and they will be attracted by the negative potential applied to the emitter and hence will flow through it to the end connected to the negative terminal of the battery. Here they will pick up an electron and disappear.

At the same time, in the base-collector circuit we will have a reverse current flowing due to the minority carriers. Holes appearing in the collector side of the depletion layer will cross the junction into the base and flow to the base terminal connected to the negative terminal of the battery, biasing the base-collector junction. Here each hole will pick up an electron and disappear. Electrons in the depletion layer on the base side of the junction will be attracted by the positive potential applied to the end of the collector. Hence they will cross the junction and flow toward the positive end of

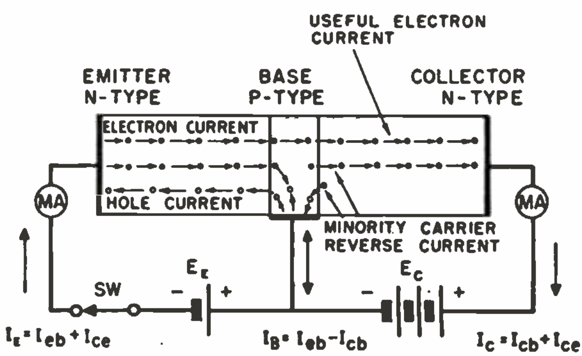


Fig. 24. Current flow and carrier movement in an NPN junction transistor.

the collector and from there to the positive terminal of the battery connected between the base and the collector.

Of these different currents flowing, the important and useful current flow is the flow of electrons from the emitter through the base to the collector. Since this is the useful current, we are interested in making this as large as possible in comparison to the other currents flowing across the emitter-base junction. Thus, the recombination current, which is due to electrons from the emitter crossing into the base and recombining with the holes, serves no useful purpose and should be kept as low as possible. This is accomplished by adding more donor atoms to the emitter than accepter atoms to the base. Thus there will be many more free electrons in the emitter than there will be holes in the base and the recombination current will be kept quite small.

Also, since there are a limited number of holes in the base compared to the number of electrons in the emitter, the number of holes crossing from the base to the emitter is also kept low in comparison to the number of electrons crossing from the emitter into the base. In a good transistor, over 95% of the electrons that cross the emitter-base junction flow to the collector.

Notice the differences and the similarities between the PNP and the NPN transistors. In both cases the emitter-base junction is forward biased and the base-collector junction is reverse biased. However, the battery connections must be reversed to provide the biases. In other words, with the PNP transistor the battery used to bias the emitter-base junction is connected with the positive terminal to the emitter and the negative terminal to the base. With the NPN transistor, the negative terminal of the battery is connected to the emitter and the positive terminal to the base. However, both are forward biased because in each case the positive terminal of the battery is connected to the P-type germanium and the negative terminal to the N-type germanium.

The base-collector junction of both transistors is reverse biased. In the PNP transistor, the positive terminal of this battery is connected to the base and the negative terminal to the collector; whereas in the NPN transistor, the negative terminal is connected to the base, and the positive terminal to the collector. Again, however, in both cases the positive terminal is connected to the N-type germanium and the negative terminal to the P-type germanium.

Also notice that in the PNP transistor the useful current flow is by means of holes, whereas in the NPN transistor the useful current flow is by means of electrons.

SELF-TEST QUESTIONS

(an) What is the base region of a transistor?

(ao) What two materials are widely used in the manufacture of transistors?

(ap) What two types of junction transistors are widely used?

(aq) What type of bias is used across the emitter-base junction in a transistor?

(ar) What type of bias is used across the base-collector junction of a transistor?

(as) Is the base region of a transistor usually a thick region or is it thin?

(at) Draw a diagram of a PNP transistor and show how the batteries are connected to place the correct bias across the two junctions.

(au) Draw a diagram of an NPN transistor and show how the batteries are connected to provide the correct bias across both junctions.

(av) What are the useful current carriers in a PNP transistor?

Semiconductor Types

There are two basic types of transistors that you will run into continuously. You are already familiar with these two types; they are the NPN transistor and the PNP transistor. However, these transistors are made in a number of different ways and the manufacturing processes result in transistors with different characteristics. In this section we are going to briefly discuss some of the important types and characteristics. We don't expect you to remember all these details; the important thing for you to remember is that they are basically either NPN or PNP transistors and operate in the same way as those we have discussed previously.

Also in this section of the lesson we'll discuss two other important semiconductor devices, the field-effect and unijunction transistors.

GROWN-JUNCTION TRANSISTORS

The first commercially available junction transistors were of the grown-junction type. This type of transistor is made from a rectangular bar cut from a germanium crystal that has been grown. Suitable impurities are added so that NPN regions such as those shown in Fig. 25 are formed. The base of the transistor is usually located midway between the two ends. Suitable contacts are then welded to the emitter, base and collector regions.

Of course, the actual bar of semiconductor material used is quite small. The emitter and the collector are considerably larger than the base; the base is kept as thin as possible and may have a thickness of less than .001".

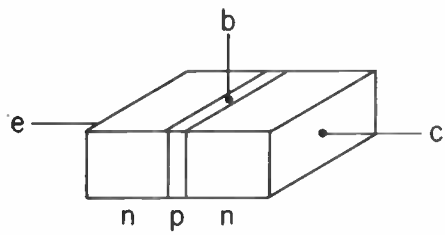


Fig. 25. A grown-junction transistor.

As mentioned the early germanium transistors were of the grown-junction type. The disadvantage of this type of transistor is that it is not particularly suitable for operations at high frequencies. In addition, it is quite temperature sensitive and can become quite unstable at higher temperatures.

ALLOY-JUNCTION TRANSISTORS

The alloy-junction transistor is made from a rectangular piece of

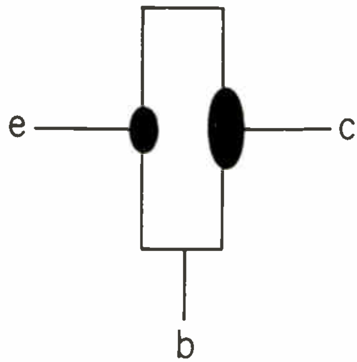


Fig. 26. An alloy-junction transistor.

semiconductor material to which suitable donor materials have been added. This results in an N-type piece of germanium or silicon. Small dots of indium are fused into the opposite sides of the wafer as shown in Fig. 26. The result is that P-type semiconductor material will be formed with the dots fused into the wafer so that we will have a PNP transistor.

An NPN-type alloy-junction transistor may be made by fusing a lead antimony alloy into each of the two opposite sides of a P-type semiconductor wafer. In this type of transistor it is possible to get a more uniform penetration of the lead antimony alloy into the semiconductor material, and this in turn leads to better junction spacing. This will cut down on the width of the space between the emitter and collector and give improved high-frequency performance. In addition, since the mobility of the electrons is more than twice that of holes, the NPN transistor will be better at high frequencies.

The general advantage of the alloy-type junction over the grown-type junction transistors is that they are usable at a somewhat higher frequency. In addition, they have a higher current gain, and the current gain remains stable as the temperature increases.

Surface-Barrier Transistor.

The surface-barrier transistor is similar to the alloy-type transistor except that depressions are etched into the N-type wafer. This permits smaller emitter and collector contacts and results in lower capacities between sections of the transistor which in turn results in better high-frequency performance.

In Fig. 27 we have shown a simplified sketch of a surface-barrier transistor. The sketch in Fig. 27B shows the carrier movement from the emitter across the base to the collector. Notice that in the sketch the emitter is shown smaller than the collector, we have shown it this way because this is the way the semiconductor is actually manufactured.

Various manufacturing techniques are used in the manufacture of the surface-barrier transistor. Both silicon and germanium types are made. In the manufacturing process different materials are evaporated or plated on to the etched depressions depending on the type of tran-

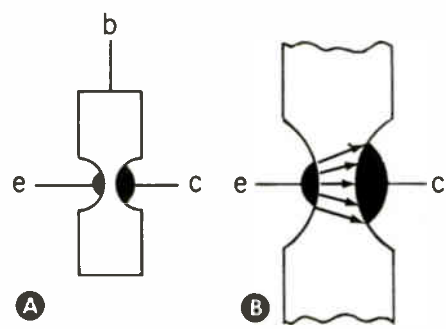


Fig. 27. Sketch of a surface barrier transistor is shown at A. Hole movement across the base is shown at B.

sistor being manufactured. However, regardless of the manufacturing technique used, which is of no interest to the technician, the surface-barrier transistors all have the characteristic of giving good performance at high frequencies.

DIFFUSION TRANSISTORS

To understand diffusion you have to understand a little about the molecular structure of materials. If you look at the wall of a glass jar, to the eye it appears solid with no space between the various molecules making up the jar. However, if you were to fill the jar with hydrogen and store it for any length of time, you would find that in a short while, the jar was no longer filled with hydrogen only, but contained a mixture of hydrogen and air. The reason is that the small hydrogen atoms are able to diffuse or pass right through the spaces between the molecules in the glass. At the same time, molecules of air will diffuse through the glass and pass on into the inside of the bottle. The hydrogen molecule is smaller than the air molecule; therefore the hydrogen will diffuse out of the jar faster than the air will diffuse in.

Diffusion can be used to add impurities to either silicon or germanium, and produce either N-type or P-type semiconductor material. The process can be controlled to provide either very uniform base, emitter, and collector regions, or it can be controlled to provide non-uniform base, emitter, and collector regions.

The Drift Type.

One of the most important uses of the diffusion technique is in the manufacture of transistors with a non-uniform base region. If the emitter and collector junctions are made by the alloy technique, but the base region is made by the diffusion technique and the impurities in the base region varied, we have what is known as a drift transistor. In a typical PNP-drift transistor, accepter impurities are added in the emitter and collector region. These impurities are controlled so that their concentration is uniform throughout the emitter and collector region. At the same time donor impurities are added to the base region. Their concentration is controlled so that it is highest in the region of the emitter-base junction and then drops off quickly and finally reaches a constant value which it maintains over to the base-collector junction, as shown in Fig. 28. This type of transistor is called a drift transistor, and its most important characteristic is its excellent performance

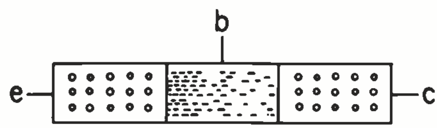


Fig. 28. Diagram showing how a large number of donor impurities increases the electron concentration in the base.

at high frequencies. However, notice that it is still a PNP transistor and the basic theory of its operation is similar to that of any other PNP transistor. The improved performance is obtained by varying the concentration of donor impurities in the base region.

The Mesa Type.

It is also possible to manufacture a transistor using the diffusion technique entirely. An example of this type is the mesa transistor.

In this type of transistor a semiconductor wafer is etched down in steps so that the base and emitter regions appear as plateaus above the collector region as shown in Fig. 29. The advantages of the mesa transistor are good high-frequency performance and very good consistency. By this we mean that it is possible to control the manufacturing techniques quite closely so that the characteristics of mesa transistors of the same type number will be quite similar. This is not necessarily true of other transistors; often their characteristics vary over a wide range.

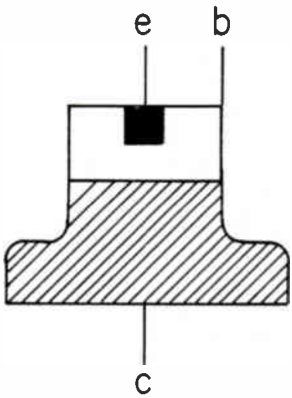


Fig. 29. A mesa transistor.

The Planar Type.

Another type of transistor manufactured by the diffusion technique is the planar type of diffused transistor. This type of transistor is shown in Fig. 30. Notice that each of the junctions is brought back to a common plane, whereas in the mesa type the various junctions are built up in plateaus. The importance of the planar-type transistor is that the junctions can be formed beneath a protective layer. As a result, many of the problems associated with other types of transistors having

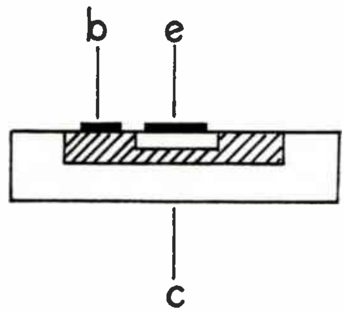


Fig. 30. A diffused planar-type transistor.

junctions exposed at the surface are avoided in this type of construction. Important characteristics of the planar transistor are generally very low reverse current and improved dc gain at low-current levels.

EPITAXIAL TRANSISTORS

One of the disadvantages of the diffusion-type transistor is the relatively high resistance of the collector region. This results in slow switching time; it limits the usefulness of the transistor in high-frequency applications. Reducing the resistance of the collector region reduces the collector breakdown voltage and this in turn again reduces the usefulness of the transistor. These problems can be overcome by the epitaxial technique. In this technique a thin high-resistance layer is produced in the collector region and the remainder of the collector region is controlled to keep its resistance low. This results in a transistor that looks something like the one shown in Fig. 31. The primary advantage of this transistor is that it provides good performance at very high frequencies. This technique can be combined with other techniques to produce transistors having varying characteristics. The epitaxial transistor can be referred to as a double-diffused epitaxial transistor. The thin high-resistance collector region is formed by the epitaxial technique and the base and the emitter are formed by the diffusion process - hence the term double diffusion.

All the transistors that we have discussed so far in this section of the lesson are either NPN or PNP transistors. The manufacturing techniques used to manufacture these transistors result in transistors of different characteristics, but the basic theory of operation of these transistors is the same. Now, we'll look at another semiconductor device which operates on a somewhat different principle.

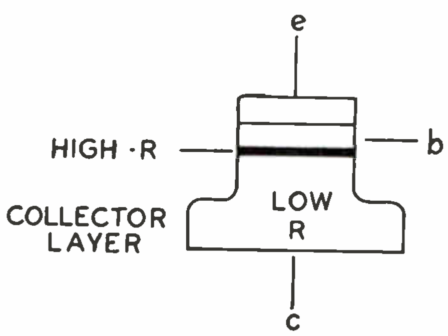


Fig. 31. A double-diffused epitaxial transistor.

THE JUNCTION FIELD-EFFECT TRANSISTOR

An interesting transistor that resembles a vacuum tube very closely in its characteristics and to some extent its operation is a field-effect transistor. One type of field-effect transistor can be made by taking a piece of N-type material as shown in Fig. 32. If the negative terminal of a battery is connected to one end of the material and the positive side

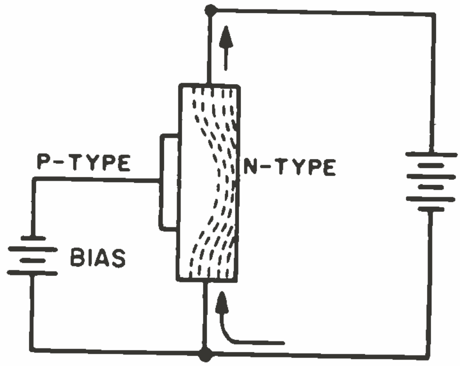


Fig. 32. Drawing showing the basic operation of a field-effect transistor.

of the terminal to the other end, electrons will flow through the material as shown. If we attach a piece of P-type material to one side so that the PN junction is formed and then place a negative voltage on the P-type material as shown in Fig. 32, there will be no current flow across the junction, because the battery biases the junction in such a way that electrons cannot flow from the N-type material to the P-type material nor can holes flow from the P-type material to the N-type.

However, the negative voltage applied to the P-type material sets up a field in the N-type material. This field opposes the electrons flowing through the N-type material and forces them to move over to one side so that the electron movement follows the path shown in Fig. 32. The negative voltage applied to the P-type material has the effect of increasing the resistance of the N-type material in the area in which the field is affected. It forms a depletion layer around the junction so there will be no free electrons in the N-type material near the junction. If the negative bias voltage is made high enough, it is able to prevent

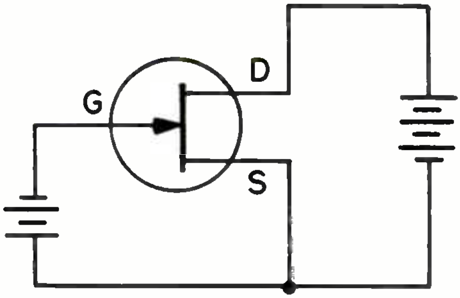


Fig. 33. Schematic representation of the circuit shown in Fig. 32.

the flow of electrons through the N-type material entirely so that the current flow will be cut off. We call this voltage where the bias voltage is high enough to stop the flow of current through the N-type material the "pinch-off" voltage. The N-type material is referred to as a channel, and the P-type material as a gate. This type of transistor is called a "junction field-effect transistor."

The schematic representation of the circuit shown in Fig. 32 is shown in Fig. 33. Notice that the end of the N-type channel at which the electrons from the battery enter is called the "source". The other end, the end from which the electrons leave and flow to the positive terminal of the battery, is called the "drain". The P-type material is called the gate, as we mentioned previously. The transistor is called a field-effect transistor because it is the field produced by the bias voltage applied to the gate that controls the flow of current through the channel. This particular type of transistor is called a junction transistor because a junction is formed between the P and N-type materials. It is called an N-channel transistor because the material in the channel through which current flows has been treated in such a way as to produce an N-type semiconductor material. Thus the complete name for this type of transistor is an N-channel, junction-gate, field-effect transistor. We usually abbreviate field-effect transistor FET, so you will see that this type of transistor is abbreviated JFET to indicate it is a junction-gate type.

An amplifier using a field-effect transistor of this type is shown in Fig. 34. In this circuit we have eliminated the bias battery by means of a resistor connected between the negative terminal of the battery and the source. This resistor might be compared to the cathode-bias re-

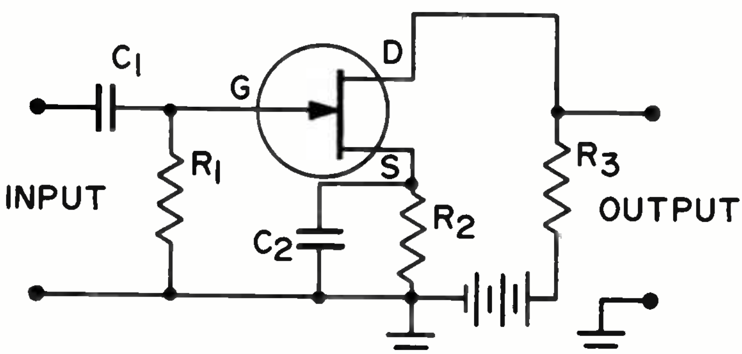


Fig. 34. An amplifier using an N-channel junction gate FET.

sistor in a triode vacuum tube amplifier stage. In the amplifier circuit, electrons flow from the negative terminal of the battery through the resistor R_2 to the source. In so doing they set up a voltage drop across R_2 having a polarity such that the source is positive with respect to ground. Since the gate connects back to ground through R_1, the gate will be at ground potential and this will make the source positive with respect to the gate, or in other words, the gate negative with respect to the source. Therefore none of the electrons in the N channel will flow to the gate, because the gate is negative.

Electrons will flow through the N channel to the drain and then through the load resistor R_3 back to the positive terminal of the battery. As the input voltage applied across the input terminals causes the voltage between the gate and the source to vary, the current flow from the source to the drain will vary because the controlling action of the gate on the current through the channel depends upon the voltage between the gate and the source. Thus we have a varying current, which will vary as the input signal varies, flowing from the source to the drain of the transistor and through the load resistor R_3. This varying current flowing through R_3 will produce an amplified signal voltage across R_3.

It is interesting to note the similarity between the circuit shown in Fig. 34 and a triode amplifier. When the input signal swings the gate in a positive direction, current flowing through the transistor will increase; this will cause the voltage drop across R_3 to increase and therefore the voltage between the drain and ground will decrease. Thus a positive-going signal applied to the gate will cause a negative-going signal at the drain. In other words, this transistor inverts the signal phase just as the triode vacuum tube amplifier stage does.

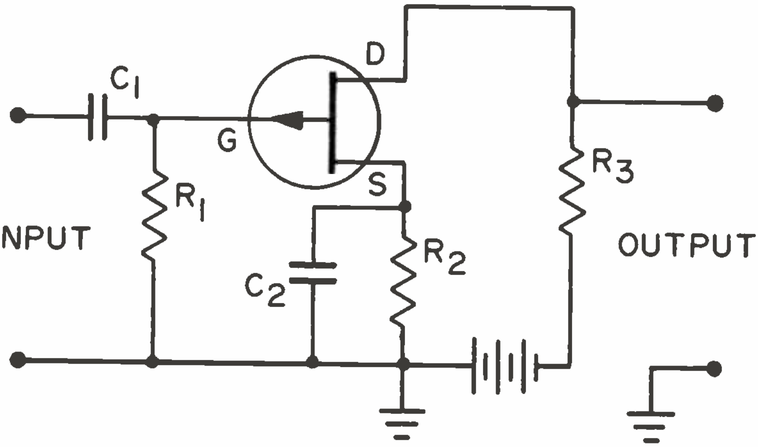


Fig. 35. An amplifier using a P-channel junction gate FET.

P-Channel JFET.

It is possible to make a P-channel junction-gate field-effect transistor by using a P-type material between the source and drain. The gate is then made of an N-type material. The bias polarity is reversed so that once again the PN junction is biased and no current flows across the junction.

A schematic diagram of an amplifier using a P-channel junction-gate effect is shown in Fig. 35. Notice the schematic symbol for the P-channel unit; we have turned the direction of the arrow around just as we did to distinguish between NPN and PNP transistors. Also notice that in this circuit the battery polarity is reversed. This is because the carriers in the channel in the P-channel unit will be holes. The positive terminal of the battery which connects to the source through R_2 repels the holes and they travel through the channel to the drain where they are attracted by the negative potential connected to the drain. Meanwhile, holes arriving at the drain terminal are filled by electrons which flow from the negative terminal of the battery through R_3 to the drain. At the same time, the positive terminal of the battery attracts electrons from the source creating new holes. These electrons flow from the source through R_2 to the positive terminal of the battery.

The operation of the P-channel, junction-gate effect is the same as with the N-channel unit, except that in one case the majority carriers are electrons, and in the other case they are holes.

In discussing the action of the junction-gate field-effect transistor, we often refer to the reverse bias across the junction creating a depletion layer in the conducting channel. In the case of an N-channel unit, the negative voltage on the P-type gate will repel electrons at the junction so that the electrons have been depleted from that area around the junction. The higher the negative voltage the further the electrons are depleted in the area around the junction, and as we pointed out previously if the voltage is made high enough, all of the electrons will be depleted so that there will be no current flow through the channel. The transistor is referred to as a depletion-type transistor because the bias depletes the number of majority carriers from the channel around the junction region. Remember what we mean by a depletion type of FET; you'll see later there is another type.

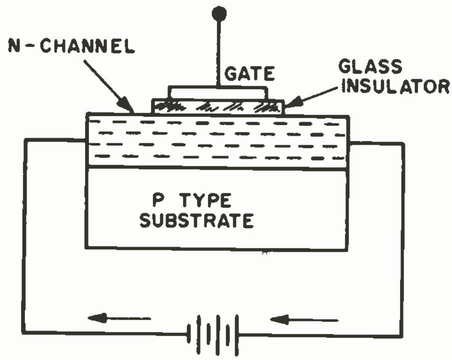


Fig. 36. Current flow through an insulated-gate, N-channel field-effect transistor with no bias applied.

INSULATED-GATE FIELD-EFFECT TRANSISTORS

The transistors we have been discussing so far are called junction-gate field-effect transistors. There is another type of field-effect transistor that is called an insulated-gate field-effect transistor. We usually abbreviate this IGFET.

In the insulated-gate field-effect transistor, the gate is completely insulated from the channel by a thin insulating material. For example, a very thin piece of glass might be placed between the conducting channel and the gate. Thus there is no actual junction formed between the semiconductor materials in the channel and the gate. In an N-channel, insulated-gate field-effect transistor, construction such as shown in Fig. 36 is often used. Here we have an N channel between the source and drain. The substrate on which the channel material is mounted is P-type material and the gate is placed along the channel as shown in the figure. The thin layer of glass prevents any actual contact between the channel and the gate.

In operation, the source and the substrate are connected to the negative terminal of the battery and the drain is connected to the positive terminal. This will permit current to flow from the negative terminal of the battery to the source, through the channel to the drain and then back to the positive terminal of the battery.

When a negative voltage is applied to the gate, it has the effect of repelling electrons away from the gate as before. In addition, the negative potential applied to the gate attracts holes in the P-type material so that the width of the channel is reduced as shown in Fig. 37. Thus the current flow through the channel is restricted by the narrowing of the channel. In effect, the resistance of the channel is increased. We refer to this type of channel as a depletion channel. The transistor is called an insulated-gate-field-effect transistor and it is also referred to as a depletion type because the flow of current through the transistor is controlled by producing a depletion layer in the channel as in the case of the junction transistors discussed previously.

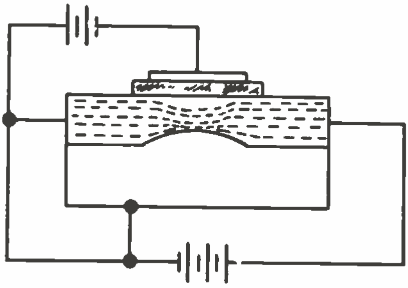


Fig. 37. Current flow through an N-channel IGFET with bias applied.

Both N-channel and P-channel IGFET's are manufactured. The schematic symbols used to represent the two different types are shown in Fig. 38A and B. In A, we have shown the symbol used for an N-channel type, and in B the schematic symbol used for a P-channel type. In operation, the units perform in essentially the same way as the junction-gate units with the exception that there will be no current

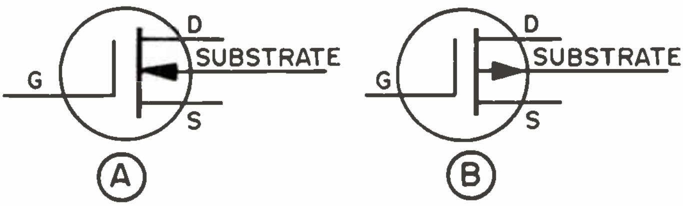


Fig. 38. Insulated-gate field-effect transistors. (A) shows the schematic symbol for an N-channel unit and (B) the symbol for a P-channel unit.

flow at all from the channel to the gate or from the gate to the channel. In the JFET, there may be very small leakage current across the junction. However, a JFET has a high input resistance because this leakage current is low. The IGFET has an even higher input resistance because there is no current flow at all from the gate to the channel or from the channel to the gate. Thus the input resistance of an IGFET is almost infinite.

Enhancement Type.

So far the field-effect transistors we have been discussing are all what are known as depletion types. In the depletion type of FET, the channel is formed and a bias is placed on the gate so as to reduce the size or width of the channel. In the enhancement-type of field-effect transistor, there is no channel present until the bias is applied to the gate. Thus, there is no current flow from the source to the drain through the transistor, unless there is a bias applied to the gate. The polarity of the bias applied to the gate is reversed from what it is in the depletion type, and this bias forms the channel through which current can flow. The operation of the units is the same as with the depletion type with the single exception of the reverse bias. In other words, in the case of an N-channel enhancement-type field-effect transistor, instead of placing a negative bias on the gate to reduce the width of the channel, as we do in the depletion-type transistor, in the enhancement-type we place a positive bias on the gate and produce the N channel.

The enhancement-type field-effect transistor is always an insulated gate type. In the case of a junction FET, if we produced an enhancement type, we would have current flow across the junction because the voltage required to produce the channel would forward bias the junction. However, in the insulated-gate FET, no current can flow across the junction because we have an insulating material between the gate and the channel. Thus we can put any type of bias we want, either forward or reverse bias, on the gate and we still will not get a current flow from the gate to the channel or from the channel to the gate.

The schematic symbol of an N-type IGFET of the enhancement type is shown in Fig. 39A. Notice that we have indicated there is no channel by breaking the channel into three parts. When the correct bias is applied to the gate, an N channel between the source and the drain will be formed. The schematic symbol for the P-channel unit of an enhancement-type IGFET is shown in Fig. 39B.

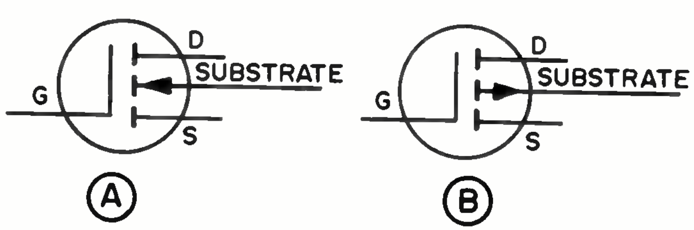


Fig. 39. A shows the schematic symbol for an N-channel enhancement-type IGFET. B shows the P-channel unit.

The operation of the enhancement-type IGFET is basically the same as with the depletion type. It could be used in a circuit similar to the circuits shown in Fig. 34 and Fig. 35.

One of the problems with IGFET's is the very high resistance between the gate and the channel. In shipping these units the manufacturer usually wraps the leads in tin foil to keep them connected together. If he doesn't do this, static charges can build up on the gate because of the very high resistance between the gate and the channel. These static charges may become high enough to actually puncture the insulation between the gate and the channel and thus ruin the unit.

In soldering an IGFET into a circuit, there might be enough leakage from the power line through the tip of your soldering iron to ruin the FET. To prevent this from happening, ground leads should be used on the various connections to the transistor and these leads should be left in place until the transistor is installed in the circuit. Once the transistor is soldered in place, you do not have to be concerned about static charges destroying the unit because the resistance in the circuit will be low enough to prevent static charges from building up to a high enough value to destroy the transistor.

Field-effect transistors are finding their way into commercial equipment, and you should therefore be sure you understand how they operate. You should review the sections on field-effect transistors several times if necessary because you can be sure they are going to be widely used in the future. They offer the advantages of the transistor as well as many of the advantages of the vacuum tube.

THE UNIJUNCTION

Another important semiconductor device is the unijunction. The unijunction is different from a conventional two-junction transistor in that it has only a single junction.

Most unijunctions are made of a bar of N-type silicon. There are two base contacts made to this bar called base 1 and base 2. These contacts are made at the ends of the bar. Between the two bases is a single rectifying contact called the emitter. The schematic symbol of the unijunction is shown in Fig. 40.

In Fig. 41 we have an equivalent circuit showing how the unijunction operates. We have referred to the resistance between base 1 and the emitter as R_{B1} and the resistance between base 2 and the emitter contact as R_{B2}. When a dc voltage is applied to the unijunction between B_1 and B_2, a current will flow through the base as shown. As long as the voltage drop across R_{B1} is greater than the emitter voltage, the emitter will be reverse biased so that there will be no current flow across the junction between the emitter and the base. The voltage across the resistance representing base 1 and the voltage across the resistance representing base 2 will remain constant. The two bases more or less act like two resistors

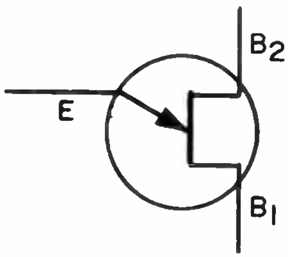


Fig. 40. Schematic symbol of a unijunction.

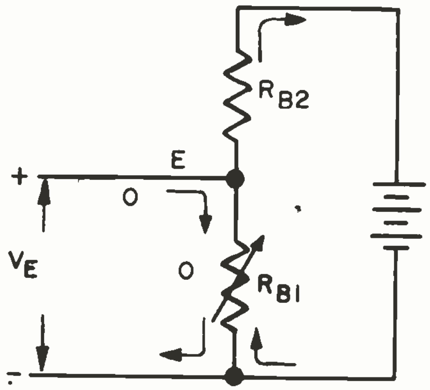


Fig. 41. Equivalent circuit showing the operation of the unijunction.

in series. The positive voltage at the emitter junction prevents any electrons from leaving the base and crossing the junction to the emitter and also prevents holes from traveling from the emitter to the base. There will be a small leakage current across the junction, but this is of no importance insofar as the operation of the unijunction is concerned.

If the voltage, V_E, exceeds the voltage across R_{B1}, then holes will enter the base and flow through R_{B1} as shown by the arrows on the diagram. These holes will cause the number of electrons flowing in R_{B1} to increase. The net result will be that you will have a drop in voltage across R_{B1} but at the same time an increase in current.

You will remember from Ohm's Law that the current flowing in a circuit is equal to the voltage divided by the resistance. If the voltage drops, the current must drop. However, in this device we have a situation where the voltage drops, but the current increases. We refer to this as "negative resistance". Devices that have this characteristic can be used in various types of amplifier circuits.

The unijunctions made for a number of years always made use of an N-type base material and a P-type emitter. However, recently some unijunctions using a P-type base material and an N-type emitter have been developed. The schematic symbol is the same, except that the direction of the arrow is reversed.

Unijunctions have not been widely used in commercial radio and TV equipment; however, they have been used in various pieces of test equipment. It is quite likely that as more transistorized television receivers are manufactured, the unijunction may be used in the sweep circuits since they are quite readily adapted to this type of application.

The important thing for you to remember at this time about the unijunction is that the device has a single junction and that the resistance of the two bases remains essentially constant until the emitter voltage exceeds the voltage across base 1. Then the voltage drop across base 1 decreases while the current flow through it increases, resulting in the negative resistance characteristic of base 1.

SUMMARY

There are too many details in this section to try to summarize them. The important thing for you to do is to realize that the different names assigned to the conventional two-junction transistors indicate the manufacturing process used to make the transistor. Typical two-junction transistors are either NPN or PNP transistors, and the basic theory of operation of the two-junction transistors is the same regardless of the manufacturing technique used. Different manufacturing techniques result in transistors with different characteristics, but the theory of operation is the same.

The field-effect transistor is a transistor that very closely resembles a vacuum tube in many of its characteristics. Remember that there are two basic types: the junction field-effect transistor and the insulated-gate field-effect transistor. In the insulated-gate type, the gate is insulated so that the leakage current to and from the gate is practically zero. This type of transistor has a very high input resistance.

You should also remember that field-effect transistors can be made in both N-channel types and P-channel types. You'll recall that by depletion type we are referring to a transistor where a channel is present. The input voltage to this type of transistor controls its channel width. JFET transistors are all of the depletion type. The IGFET may be either the depletion type or the enhancement type.

The unijunction is a semiconductor device with a single junction. Its use in commercial equipment is somewhat limited at this time, but you should understand the basic fundamentals of the device because it is quite likely that it will be used in the future.

One important point about all types of transistors that we must emphasize is that they are all easily damaged by excessive heat. This is true particularly of germanium transistors, but silicon transistors can also be destroyed by excessive heat. Whenever you have to replace a transistor in a circuit, you should make sure that the point at which you have to solder the transistor in the circuit is clean so that the solder will melt and flow over the connection quickly. Also make sure that the transistor leads are clean. It is a good idea to use a heat sink between the point at which you are soldering and the semiconductor device. A good heat sink is a pair of longnose pliers; simply hold the lead securely in the jaws of the pliers while you are soldering the lead in place. Much of the heat developed at the joint will flow through the pliers and keep the semiconductor device itself from becoming excessively hot. The joint should be soldered as quickly as possible; get the iron off the joint just as soon as the solder has melted and flowed smoothly over the connection.

Semiconductor devices can be damaged by storing them in excessively warm places. Again, this is particularly true of germanium transistors which are more heat sensitive than silicon transistors. Storing semiconductor devices at room temperature will prevent this type of damage. You should avoid storing them in any place where they can become excessively hot.

Now to check yourself on this important section you should answer the following self-test questions.

SELF-TEST QUESTIONS

(aw) Into what two basic types can the grown-junction transistor be divided?

(ax) What type of transistor can the surface-barrier transistor be classified as?

(ay) What is the most important characteristic of the surface-barrier transistor?

(az) What do we mean by a diffussion transistor?

(ba) What is an important use of the diffusion technique in manufacturing transistors?

(bb) What is the difference between a junction-gate field-effect transistor and an insulated-gate field-effect transistor?

(bc) What is a depletion-type field-effect transistor?

(bd) What is an enhancement-type FET?

(be) What is a unijunction?

Answers to Self-Test Questions

(a) Four.

(b) Germanium and silicon.

(c) A covalent bond is the sharing of two electrons by two atoms, one from each of the two atoms.

(d) Four. A single atom of germanium or silicon will share an electron from its outer ring and an electron from the outer ring of a nearby atom to form a covalent bond. It will do this with four electrons to establish four covalent bonds.

(e) Intrinsic conduction is conduction due to the formation of hole-electron pairs throughout a germanium or silicon crystal.

(f) No.

(g) Germanium.

(h) Heat.

(i) Silicon.

(j) An N-type material is a material that has been doped so that electrons are the majority carriers. This is brought about by using an impurity that has five electrons in the valence ring so that when it forms covalent bonds with nearby germanium or silicon atoms there will be an electron left over.

(k) A donor material is an impurity which when added to silicon or germanium will form covalent bonds with four nearby atoms and have an electron left over. When a donor material is added to germanium or silicon, N-type material is formed.

(l) Arsenic and antimony.

(m) P-type semiconductor material is a material that has been doped with an impurity having three electrons in the valence ring. This will leave a covalent bond that is short one electron so there will be a hole in the bond. The hole is in effect a positive charge and hence the majority carriers in the P-type material are the holes or positive charges.

(n) An accepter-type impurity is an impurity with three electrons in the valence ring or shell. It is an accepter-type material because it leaves a hole in the covalent bond which can accept an electron.

(o) Indium, boron and aluminum.

(p) Electrons are the majority carriers in N-type material.

(q) Holes.

(r) When the arsenic loses an electron it will be short one electron to completely neutralize the charge on the nucleus, and therefore the atom will have a positive charge. Meanwhile the atom of silicon or germanium that has received the extra electron will have a negative charge on it.

(s) There is no charge on the crystal, it is neutral. Although some regions may have a positive charge, other regions may have a negative charge; the crystal itself neither gains nor loses electrons and therefore it does not have any charge.

(t) Holes are produced.

(u) Diffusion is a random motion of the carriers in a semiconductor material. It goes on at all times in the crystal and every effort is made to keep diffusion as low as possible since it contributes nothing insofar as the usefulness of the material in semiconductor devices is concerned.

(v) Drift.

(w) Electrons are the majority carriers in an N-type material and they move from the end to which the negative potential is applied towards the end to which the positive potential is applied.

(x) Holes are the majority carriers in a P-type material and they move from the end to which the positive potential is applied to the end to which the negative potential is applied.

(y) No - the crystal will remain electrically neutral. In the case of N-type material, exactly the same number of electrons will leave the positive end of the crystal and enter the negative end of the crystal. In the case of the P-type material, electrons will leave the end to which the positive potential is connected creating holes. Exactly the same number of electrons will enter the end to which the negative potential is connected to fill holes arriving at the negative end.

(z) No. For a given potential and given size of crystal, electrons will move at approximately twice the rate through an N-type crystal as the holes will through a P-type crystal.

(aa) The N-type material will have the lower resistance. This is due to the higher mobility of the electrons in the N-type material than the holes in the P-type material.

(ab) Detectors and rectifiers.

(ac) The depletion layer is an area on both sides of the junction. On the P-side of the junction there is a shortage of holes and on the N-side of the junction there is a shortage of electrons. The shortage is caused by a few of the majority carriers crossing the junction in each way building up a charge at the junction so that the majority carriers are repelled away from the junction.

(ad) The potential barrier is the voltage built up across the junction by the diffusion of majority carriers across the junction. The holes that diffuse across the junction into the N-side of the junction create an area that has a negative charge in the P-side of the junction. Similarly, the electrons diffusing across the junction into the P-side create an area on the N-side of the junction that has a positive charge. This charge across the junction eventually becomes high enough to prevent any further diffusion of holes and electrons across the junction.

(ae) No. The net charge on the crystal will remain zero. There may be areas on the crystal that have a positive charge, and other areas that have a negative charge, but since the crystal itself neither gains nor loses electrons, the net charge on the crystal will remain zero.

(af) Yes. Minority carriers crossing the junction tend to weaken the potential barrier established across the junction by majority carriers diffusing across the junction. When the potential barrier is weakened, additional majority carriers can cross the junction. Thus we end up with carriers crossing the junction in both directions. This adds nothing to the useful current that the diode can handle, but it does contribute to heating and thus limits the useful current that can cross the junction.

(ag) When a junction is forward biased we have a positive potential applied to the P-side and a negative potential applied to the N-side. This permits electrons to freely cross the junction from the N region to the P region. Similarly holes can cross the junction from the P region to the N region.

(ah) When a junction is reverse biased we have a negative potential connected to the P region and a positive potential connected to the N region. The positive potential connected to the N region repels holes in the P region away from the junction so that they cannot cross the junction. Similarly, the negative potential applied to the P region repels electrons in the N region away from the junction so that they cannot cross the junction. When a junction is reverse biased, majority carriers normally cannot cross the junction.

(ai) When there is no voltage applied to a semiconductor diode, the net current flow across the junction is zero. However, in the case of a vacuum tube where there is no voltage applied between plate and cathode, some electrons will leave the cathode with sufficient energy to travel over to the plate. As a result, there will be a small current through the diode even though there is no voltage applied between the plate and cathode.

(aj) When a semiconductor diode is reverse biased, there will be a small current flow across the junction due to minority carriers. As long as the breakdown voltage of the diode is not exceeded, this current will be quite small. In the case of a vacuum tube, when the plate is made negative with respect to the cathode, the plate will repel electrons so that there will be no current flow through the vacuum tube.

(ak) A Zener diode is a diode used in applications where a reverse bias is placed across the junction. The diode is designed to break down at a certain voltage and then maintain a constant voltage. If the voltage tries to increase above this constant value, the current flow through the Zener diode will increase so that the diode can be used in voltage regulating circuits and also can be used as a voltage reference source.

(al) A tunnel diode is a diode where the electrons cross the junction by a process similar to tunneling across the junction. The tunnel diode has a characteristic of introducing negative resistance into the circuit when a certain voltage is applied across the junction. In other words, when the voltage across the diode increases, the current flow through the diode decreases. Similarly, when the voltage decreases the current increases. Because of this negative resistance characteristic, the tunnel diode can be used as an oscillator.

(am) A p-i-n diode is a diode that is primarily used as a variable resistance. The resistance of the diode varies as the voltage across it is varied. The p-i-n diode is used in automatic gain control circuits to vary the strength of the signal reaching amplifier stages.

(an) The base region is the center region of the transistor. On one side of the base region is the emitter, and on the other side is the collector.

(ao) Germanium and silicon.

(ap) PNP transistors and NPN transistors.

(aq) Forward bias.

(ar) Reverse bias.

(as) The base region is usually comparatively thin.

(at) See Fig. 22.

(au) See Fig. 24.

(av) Holes are useful current carriers in a PNP transistor.

(aw) NPN and PNP transistors.

(ax) An alloy-type transistor.

(ay) Good high-frequency performance.

(az) A diffusion transistor is a transistor which has been made by diffusing the impurities into the emitter, base and collector regions.

(ba) One of the most important uses of the diffusion technique is in the manufacture of non-uniform base regions.

(bb) In a junction-gate field-effect transistor there is an actual contact between the channel material and the gate. There will be some current flow across the contact at all times due to minority carriers crossing the junction. In addition, if the junction is forward biased there will be a high current flow across the junction. In an insulated-gate field-effect transistor a glass or similar insulating material is used between the material in the channel and the gate. Since there is an insulator between the gate and the channel, there will be little or no current flow across the insulator either due to minority carriers when there is a reverse bias applied, or due to majority carriers with a forward bias applied.

(bc) A depletion-type FET is a unit in which the channel is present at all times. The transistor works by depleting or reducing the size of the channel.

(bd) An enhancement FET is a unit in which there is no channel present until the operating bias is applied between the gate and the material in which the channel is formed.

(be) A unijunction is a semiconductor device having two base connections but only a single junction. The junction is called the emitter. The single junction makes the unijunction quite different from the conventional two-junction transistor.

LESSON QUESTIONS

Be sure to number your Answer Sheet B112.

Place your Student Number on every Answer Sheet.

Most students want to know their grades as soon as possible, so they mail in their answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable. However, don't hold your answers too long; you may lose them. Don't hold answers to more than two sets of lessons at any time, or you may run out of lessons before new ones can reach you.

1. Name the two most important semiconductor materials used for transistors.

2. When a donor type of material is added to a silicon or a germanium crystal, what type of semiconductor material is produced? Does this type have free electrons or free holes?

3. What effect do the two layers of ionized atoms at the junction in a PN diode have on the majority carriers in the vicinity of the junction?

4. To which side of a PN junction diode do you connect the positive battery terminal if you wish to place a forward bias on the junction?

5. If a reverse bias is applied to a junction diode, what effect will a small increase in bias have on the current flowing, provided the reverse voltage does not exceed the breakdown voltage?

6. What is a Zener diode?

7. In a PNP transistor, what happens to a hole that crosses the emitter and the base and moves into the collector?

8. What is a drift transistor?

9. What is an N-channel, junction-type field-effect transistor?

10. What do we mean when we refer to a field-effect transistor as an enhancement type?

CASHING IN ON DISCONTENT

Discontent is a good thing--if it makes you want to do something worthwhile. If you had not been discontented, you would never have enrolled for the NRI course.

Practically everyone is discontented. But some of us are "floored" by discontent. We develop into complainers. We find fault with anything and everything. We end up as sour and dismal failures.

Those of us who are wise use our discontent as fuel for endeavor. We keep striving toward a goal we have set for ourselves. We are happy in our work. We face defeat, and we come out the victors.

At this minute you may be discontented with many things--your progress with your course, your earning ability, yourself.

Make that discontent pay you dividends. Don't let it throw you down. If you do, you may never be able to get up again. Keep striving to remove the cause of your discontent. Remember that it's always darkest before the dawn. And a real NRI man works hardest and accomplishes most when he is face to face with the greatest discouragements.

J. M. Smith

nri

ACHIEVEMENT THROUGH ELECTRONICS

HOW TRANSISTORS ARE USED

B113

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

HOW TRANSISTORS ARE USED

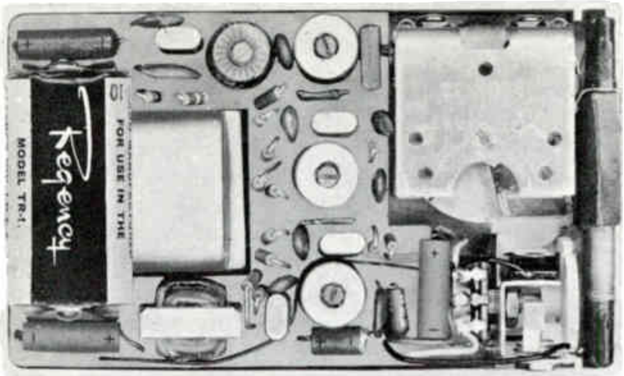
B113

STUDY SCHEDULE

By dividing your study into the steps given below, you can get the most out of this part of your NRI Course in the shortest possible time. Check off each step when you finish it.

NATIONAL RADIO INSTITUTE, WASHINGTON, D.C. 20016

P113M1074 **1975 EDITION** **Litho in U.S.A.**



HOW TRANSISTORS ARE USED

Just a few years ago a germanium transistor cost considerably more than most vacuum tubes, and the cost of silicon transistors was many times that of germanium transistors. Silicon transistors were priced so high that their use in commercial entertainment devices was prohibitive. Even germanium transistors were priced so high that it was hard to justify their use in entertainment-type equipment. Now, however, both germanium and silicon transistors are relatively inexpensive. There are many transistors of both types priced lower than even the most inexpensive vacuum tubes. For this reason they are widely used in commercial equipment. All portable radio receivers manufactured today use transistors. You will not run into a portable receiver using tubes except in the case of a receiver that is many years old. Automobile radios use transistors, and most stereo and hi-fi equipment is made entirely with solid-state devices. There are many portable television receivers on the market and in use today that use transistors exclusively. In addition, there are hybrid receivers that use both tubes and transistors.

TRANSISTORS AND TUBES

A transistor can do almost anything a vacuum tube can do. Sometimes a transistor can perform a task better than a vacuum tube; sometimes it cannot perform the task as well as a vacuum tube.

Advantages of Tubes.

The big advantage vacuum tubes have over transistors is that it is usually possible to obtain a higher gain with a single vacuum tube than it is with a single transistor. In addition, a number of tubes can be combined in a single envelope so that a single vacuum tube may actually contain three or more separate tubes in the same envelope. This often results in the multi-purpose tube being more economical than the equivalent number of transistors that would be required to perform the same functions.

Another advantage of the tube is that today's engineers are more used to working with them. Thus the cost of designing a piece of electronic equipment, such as a color TV receiver, using tubes is less than the cost of designing a color TV receiver using transistors.

The characteristics of tubes are more uniform than those of transistors. In other words, it is easier to duplicate circuits using tubes than circuits using transistors. Furthermore, when it comes time to replace a transistor, you may find the replacement has quite different characteristics from the original and that the stage will perform quite differently from the manner in which it did with the original transistor. You are not likely to run into this situation with a tube.

Advantages of Transistors.

To offset the advantages of the vacuum tube, the transistor offers many other advantages. Generally speaking, a transistor for a particular job is smaller than a vacuum tube. The transistor is more rugged than the vacuum tube and there is less chance of it breaking. A transistor does not require any power to heat a cathode or filament - this reduces the power requirement of the equipment and also reduces the amount of heat that must be dissipated by the equipment. The lower temperature inside the equipment generally results in longer life from all the other parts in the equipment.

Transistors operate on lower voltages than vacuum tubes. This often results in savings in the power supply and also in the voltage ratings of the other components in the circuit.

BASIC TRANSISTOR CIRCUITS

In the span of the few short years in which transistors have been manufactured for commercial applications, there have been a large number of different transistor types manufactured. Each year manufacturers introduce new types, and we can expect that this will continue year after year. This might make you think that it will be an almost impossible job to keep up with new developments in the field of transistors. However, this is not the case. Like tubes, transistors are used in certain basic circuits. If you learn these basic circuits and how they work, you should be able to understand new circuits as you encounter them. In addition, there are certain basic characteristics of transistors that are important. Once you learn what these characteristics are and what they mean, you will be in a position to evaluate new transistors as they appear on the market in comparison with older transistors with which you may be more familiar.

As you will remember from your study of vacuum tubes, the triode tube has three elements: a cathode, a plate and a grid. You will remember that we found that there are three different types of circuits in which a triode tube can be used: the grounded-cathode circuit, the grounded-plate circuit, and the grounded-grid circuit. Similarly, a triode transistor has three elements: an emitter, a base, and a collector. There are three basic circuits in which a triode transistor can be used. These circuits are called the common-base circuit, the common-emitter circuit, and the common-collector circuit. In transistor circuits we usually use the word "common" when speaking of different circuits, but actually a common-base circuit is a grounded-base circuit--in other words a circuit where the base is at the ac ground potential.

Now let's study the three basic circuits to learn something about the important characteristics of each type. It is important for you to understand these three circuits. If you understand them you should have no difficulty with the other circuits you encounter in this book because they will all be variations of one of these three circuits.

Of the three circuit variations in which a triode transistor may be used, the common-emitter circuit is found more frequently than the other two. This is the circuit that will give the greatest voltage and power gain. However, we will start our study of the three circuits with the common-base because you have already seen the circuit in the preceding lesson and also because it is a little easier to understand than the other two circuits.

In studying these three basic circuits we will compare their characteristics with the three basic vacuum-tube circuits. You will see a great deal of similarity between the three circuits insofar as performance is concerned, but you will also see some very noticeable differences. You should keep in mind that, although the end results may be the same, there is a great deal of difference between the way a circuit using a vacuum tube and one using a transistor works.

The Common-Base Circuit

Typical common-base circuits are shown in Fig. 1. The circuit shown at A is for an NPN transistor; the circuit shown at B is for a PNP transistor.

The solid arrows on the two diagrams indicate the direction of useful electron flow. In the circuit shown at A, electrons flow from the negative terminal of battery B1, through resistor R1, through the NPN transistor, into the emitter, across the emitter-base junction to the base, then across the base-collector junction into the collector, and from the collector through the collector resistor R_C to the positive terminal of battery B2. Notice that the emitter-base junction is forward biased, and the base-collector junction is reverse biased as in all transistor circuits.

In the circuit shown in Fig. 1B, the batteries are reversed. Since this is a PNP transistor, in order to place a forward bias on the emitter-base junction, the positive terminal of the battery must be connected to the emitter and the negative terminal to the base; similarly, battery B2 is reversed because, to place a reverse bias on the base-collector junction, the negative terminal must be connected to the collector and the positive terminal to the base.

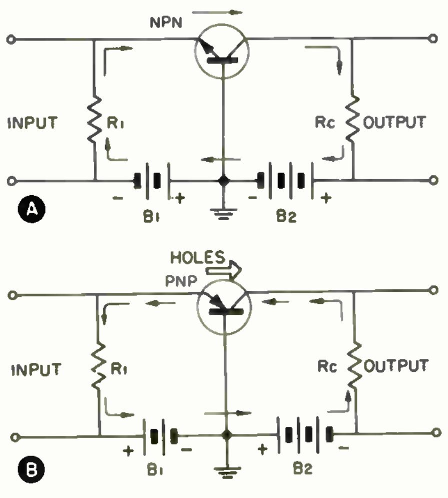


Fig. 1. Basic common-base circuits. The circuit at (A) is for an NPN transistor; the one at (B) for a PNP transistor.

In Fig. 1B the positive terminal of B1 is connected to the emitter through resistor R1. The positive potential applied to the emitter will attract electrons from the emitter. When an electron is attracted from the emitter it will flow in the direction indicated by the solid arrows, through resistor R1 to the positive terminal of B1. Meanwhile, when an electron is pulled from the emitter, a hole is created. The hole travels through the transistor in the direction indicated by the outlined arrow. The hole crosses the emitter-base junction, then travels through the base, across the base-collector junction, and to the terminal of the collector that is connected to the collector resistor R_C. There the hole is filled by an electron and disappears. The electrons needed to fill the holes reaching the collector terminal are supplied by battery B2. Thus there will be an electron flow from the negative terminal of this battery through the resistor Rc to the collector terminal of the transistor as shown by the solid arrows on the diagram.

You will remember from the preceding lesson that the majority carriers diffusing across the junctions in a transistor set up potential barriers which prevent an additional flow of majority carriers across the junction. In the transistor circuits shown in Fig. 1, battery B1 places a forward bias on the emitter-base junction that partially overcomes this potential barrier and allows some majority carriers to cross the junction. The exact number of majority carriers that will cross the junction depends upon the characteristics of the transistor and upon the voltage of battery B1. Thus, in the circuits shown in Fig. 1, we will have a static current flowing. Static current is simply a fixed current or a current flow that depends upon the operating voltages applied and not upon the signal voltage. This current is often called the "zero-signal" current, and it will set up a voltage drop across the collector resistor R_C.

A COMMON-BASE NPN AMPLIFIER

Now let us see what happens when a signal voltage is applied to the input of these transistor circuits. In Fig. 2 we have shown a common-base amplifier circuit using an NPN transistor. Let's see how this circuit can be used to amplify a signal.

Let us consider first the voltage across R_C. The end of R_C that is connected to battery B2 is essentially at signal ground potential. This is because B2 is a low impedance at signal frequencies or is made to act like a low impedance by shunting it with a capacitor. So let's consider this end of R_C as being at ground or zero potential and see what the voltage is at the other end.

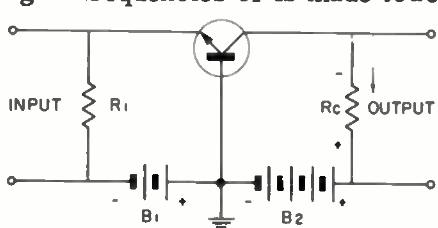


Fig. 2. A common-base amplifier circuit using an NPN transistor.

Since current flows through R_C in the direction shown, the end of R_C connected to the collector is negative. We can show this on a graph as in Fig. 3A. The voltage is negative so we represent it by a line drawn below the zero-voltage axis.

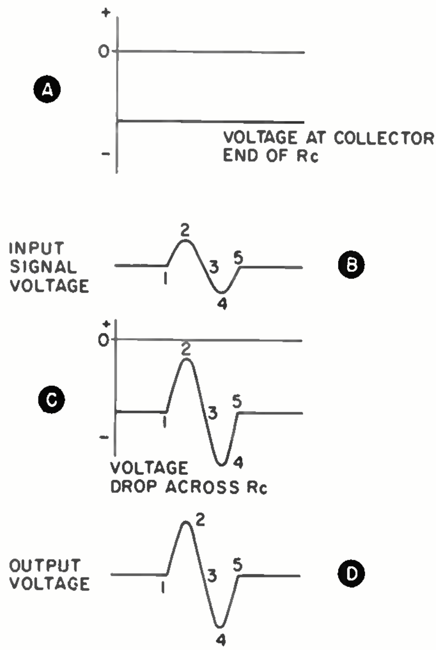


Fig. 3. Voltage waveforms for the circuit of Fig. 2.

Now, what happens when we apply a signal like Fig. 3B to the input? Consider first the input signal at point 1 in Fig. 3B. At point 1 the signal is zero and hence it has no effect on the static or zero-signal current flowing through the transistor. The only current flowing will be the zero-signal current due to the battery voltages. Therefore the voltage across R_C can be represented by the straight line extending to point 1 in Fig. 3C. Now as the ac signal moves to point 2 on the input curve, we have a voltage drop across resistor R1. The polarity of this voltage drop makes the end of R1 connected to the emitter positive and the end connected to the battery negative. This means that the polarity is opposite to the polarity of B1. Therefore the voltage across R1 will subtract from the voltage of battery B1 insofar as the net emitter-base voltage is concerned. This means that the signal will reduce the forward bias applied between the emitter and base and hence reduce the number of majority carriers (electrons) that can cross the emitter-base junction. When the number of carriers crossing this junction decreases, the number of electrons flowing through R_C will decrease.

When the number of electrons flowing through R_C decreases, the voltage drop across the resistor will decrease. This is shown in Fig. 3C. The curve rises, gets closer to zero, between points 1 and 2. This shows that the voltage is decreasing.

When the input voltage drops to point 3, we once again have the situation where the input voltage is zero. The current flowing through the collector resistor R_C will increase to the zero signal current, and the output voltage shown in Fig. 3C will increase to the zero signal voltage, point 3, which is the voltage that appeared across this resistor before any signal was applied. Hence the current will increase.

Now let us see what happens when the input signal swings in the opposite direction. When the signal swings to point 4, the end of resistor R1 that is connected to the emitter will be negative and the end connected to the battery will be positive. Now we have a voltage across R1 that is in series with the voltage of battery B1 and hence adds to it. This means that the forward bias across the emitter-base junction will be increased and the number of majority carriers crossing the junction will increase. The current flow through R_C will increase. When current flow through this resistor increases, the voltage drop across the resistor will increase and the end of the resistor that is connected to the collector will become more negative with respect to the other end. Hence the voltage appearing across the output terminals will swing to point 4 as shown in Fig. 3C. When the input signal drops back to point 5 or zero signal voltage, the voltage across the output similarly will fall back to zero signal voltage at point 5 on the output voltage curve.

Now let's consider what is happening in the output of this amplifier. First we have a dc voltage across R_C; this is the static or zero signal voltage. When a signal is applied to the input, the zero signal voltage varies. If we remove the zero signal voltage, which we can easily do by taking the output off through a capacitor, we have the output voltage shown in Fig. 3D. This is the actual amplified output obtained from the transistor.

Two important things to notice in this circuit are that the output voltage is in phase with the input voltage and that the output voltage is several times the input voltage. In other words when the input voltage swings positive, the output voltage goes positive and when the input voltage goes negative, the output voltage goes negative. We can obtain a voltage gain using this type of circuit.

A COMMON-BASE PNP AMPLIFIER

When a PNP transistor is used as an amplifier in a common-base circuit as shown in Fig. 4, the output voltage is also in phase with the input voltage. However, the way in which this circuit operates is quite different from the way the NPN circuit operates.

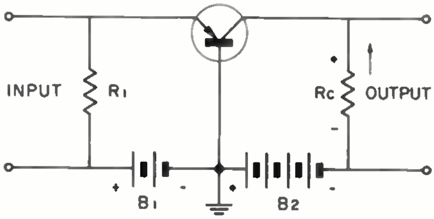


Fig. 4. A common-base amplifier circuit using a PNP transistor.

Notice that the polarity of the voltage across R_C is the opposite to what it was in Fig. 2. In Fig. 4, the end of the resistor connected to the collector is positive and the end connected to battery B2 is negative. Thus the voltage at the collector end of R_C will be positive with respect to the other end of the resistor, and under zero signal conditions can be represented by a straight line above the zero axis as shown in Fig. 5A.

Now, what happens when an input signal like Fig. 5B is applied to this circuit? When the input voltage is zero, the collector current flowing through R_C will be the zero signal current as shown in Fig. 5C extending to point 1 on the curve. But when the voltage applied across R1 swings positive to point 2 in Fig. 5B, the emitter end of this resistor is positive, and the end connected to the battery will be negative. This means that the voltage will be in series with the voltage of battery B1 and hence will increase the forward bias on the emitter-base junction. This will cause an increased movement of holes through the transistor and hence an increase in current flow through R_C. The increase in current flow through R_C results in an increased voltage drop across R_C so that the end of the resistor connected to the collector becomes more positive with respect to the other end. Thus when the input voltage moves from point 1 to point 2 in Fig. 5B, the voltage across R_C will move from point 1 to point 2 as shown in Fig. 5C.

Similarly, when the input voltage swings negative, the voltage across R1 will subtract from the forward bias applied across the emitter-base junction, reducing the bias. This will reduce the hole movement

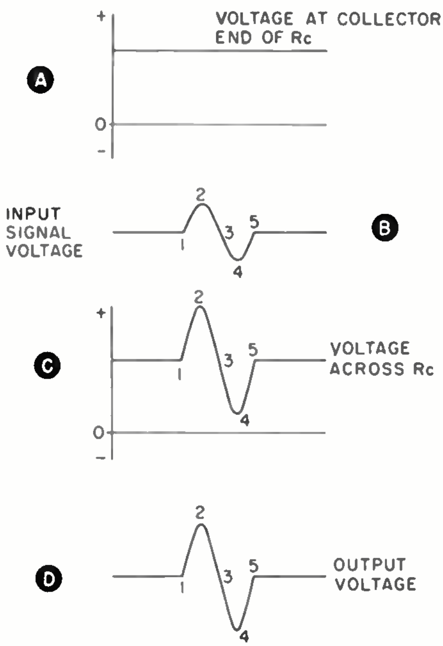


Fig. 5. Voltage waveforms for the circuit of Fig. 4.

through the transistor, which will reduce the number of electrons flowing through R_C. When the number of electrons flowing through this resistor decreases, the voltage drop across the resistor decreases with the result that the end of the resistor connected to the collector will become less positive with respect to the other end. This means that when the input signal swings from point 3 to point 4 on the input curve as shown in Fig. 5B, the signal voltage appearing across R_C will swing from point 3 to point 4 as shown in Fig. 5C.

If we once again remove the zero signal current from our graph, we have the graph shown in Fig. 5D. This represents the actual output signal voltage and, as you can readily see, it is in phase with the input signal.

CHARACTERISTICS OF COMMON-BASE CIRCUITS

Even though the action of the PNP transistor is quite different from the action of the NPN transistor, the net result using the common-base circuit is the same with both types of transistors. In both cases we have the output voltage in phase with the input voltage, and we have voltage amplification. In other words, the output signal voltage is greater than the input signal voltage.

You will remember that when you studied this circuit in the preceding lesson you learned that not all majority carriers leaving the emitter and crossing the emitter-base junction will reach the collector. Some of these carriers will be attracted by the potential of the emitter-base battery and flow out of the base to the battery. Therefore, since all of the emitter current does not reach the collector, the collector current will be less than the emitter current. Technicians say that the current gain is less than 1. For example, if the current in the emitter circuit increases by 1 milliampere, the current in the collector circuit will increase, but the increase will be something less than 1 milliampere.

Two other characteristics of a transistor amplifier that are important are the input impedance and the output impedance. The input impedance is simply the ratio of the signal voltage over the signal current. If we represent the signal voltage by e_{IN} and the signal current by i_{IN} and the impedance by Z_{IN} then the input impedance will be:

$$Z_{IN} = \frac{e_{IN}}{i_{IN}}$$

The output impedance is the ratio of the output signal voltage over the output signal current. If we represent the output signal voltage by e_{OUT} and the output current by i_{OUT} and the output impedance by Z_{OUT}, then the output impedance will be:

$$Z_{OUT} = \frac{e_{OUT}}{i_{OUT}}$$

If we examine the common-base circuit shown in Fig. 2, we see that the input voltage is applied across R1. This will cause some signal current to flow through the resistor R1. In addition, the entire emitter current drawn by the transistor must flow through R1. Therefore, even with the small signal voltage the signal current must be quite high. This means that the ratio of the voltage divided by the current will be low or, in other words, we will have a low input impedance.

On the other hand, since the stage is capable of giving voltage gain, the output voltage which will be developed across R_C will be much higher than the input voltage. At the same time, since the collector current is less than the emitter current the signal current flowing in the output will be lower than the signal current flowing in the input. Therefore the output impedance will be considerably higher than the input impedance. In fact, the output impedance will be quite high. Therefore in the common-base amplifier we have a very low input impedance and a reasonably high output impedance.

COMMON-BASE AND GROUNDED-GRID CIRCUITS COMPARED

The common-base circuit is often compared to the grounded-grid vacuum tube circuit. The two circuits are shown in Fig. 6 for comparison purposes. Notice the similarity between the common-base circuit

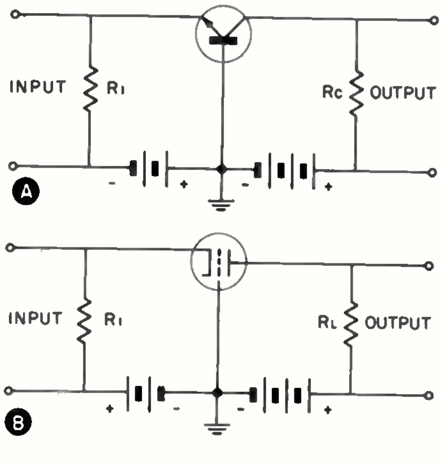


Fig. 6. Comparison of a common-base transistor circuit (A) with a grounded-grid vacuum tube circuit (B).

shown at A and the grounded-grid circuit shown at B. The battery shown between the grid and the cathode of the grounded-grid circuit is seldom found in practice, because resistor R1 can be made to supply the bias required by the tube or, in some instances, a resistor placed in the grid circuit is used to develop bias. In this case the grid resistor will be bypassed by a capacitor.

SUMMARY

The common-base circuit has a very low input impedance and a high output impedance. This output voltage is in phase with the input voltage and is greater than the input voltage. The current gain of the stage is less than one. A common-base circuit is often used in applications where we want to match a low impedance to a high impedance.

SELF-TEST QUESTIONS

(a) In the common-base circuit, is it possible to get a current gain?
(b) What is the phase relationship between the amplified signal voltage and the input signal voltage in a common-base circuit?
(c) What are the majority carriers in a common-base circuit using a PNP transistor?
(d) Why is it possible to get a voltage gain in a common-base amplifier circuit even though we do not have a current gain?
(e) What are the relative input and output impedances of the common-base amplifier circuit?
(f) To what type of vacuum tube circuit can the common-base circuit be compared?

The Common-Emitter Circuit

The most frequently used transistor circuits are the common-emitter circuits shown in Fig. 7. The circuit shown at A is for an NPN transistor, and the circuit shown at B is for a PNP transistor. Battery B1 in both cases provides the forward bias needed for the emitter-base junction, and battery B2 provides the reverse bias needed for the base-collector junction. In Fig. 1A battery B1 was connected so that the emitter was made negative with respect to the base. In Fig. 7A the emitter is also made negative with respect to the base. This provides forward bias for the emitter-base junction of the NPN transistor.

The solid arrows on the two diagrams in Fig. 7 indicate the direction of electron flow in the circuit. In Fig. 7A electrons leave the negative terminal of B1 and flow to the emitter. They cross the emitter-base junction, then flow through the base, across the base-collector junction, and through resistor R_C to the positive terminal of B2.

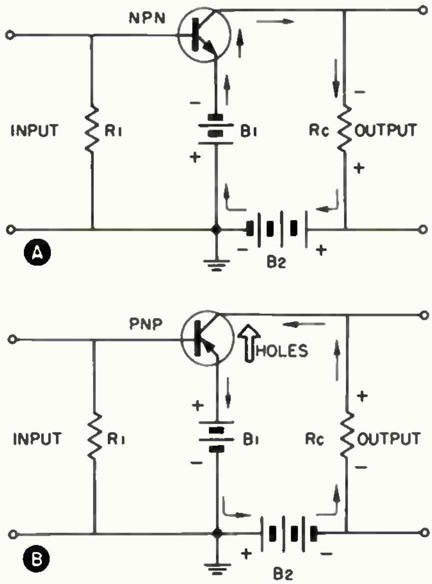


Fig. 7. Common-emitter circuits for (A) an NPN transistor, and (B) a PNP transistor.

In the circuit shown in Fig. 7B the electrons are attracted from the emitter by the positive potential of B1 and flow to the positive terminal of B1. The electrons leaving the emitter leave holes behind. These holes flow through the emitter, across the emitter-base junction, through the base, across the base-collector junction, and to the terminal of the collector. Here the holes are filled by electrons supplied by battery B2. The electrons from B2 leave the negative terminal of the battery, flow through the collector resistor R_C and to the terminal of the collector.

Notice the difference between these circuits and the common-base circuits shown in Fig. 1. In the common-base circuits, the useful transistor current flows through the input resistor R1, whereas in the circuits shown in Fig. 7, the useful transistor current does not flow through R1. The result is that the common-emitter circuit has a much higher input resistance than the common-base circuit. This is an advantage because it means that the generator driving the common-emitter circuit does not have to have such a low output impedance.

Notice that in the two circuits shown in Fig. 7 electrons flow through the collector resistor R_C in opposite directions. In the circuit shown at A, the electrons flow from the collector through the resistor to the battery, making the end of the resistor connected to the collector negative with respect to the other end. In the circuit shown at B, electrons flow from the negative terminal of B2 to the collector, making the end of the resistor connected to the battery negative with respect to the end connected to the collector. In other words, the polarity of the voltage across R_C in the circuit shown in Fig. 7A is opposite to the polarity of the voltage across R_C in the circuit shown in Fig. 7B. Now let us see how this type of circuit works.

A COMMON-EMITTER NPN AMPLIFIER

A common-emitter circuit using an NPN transistor is shown in Fig. 7A. As in the common-base circuit you just studied, battery B2 is a low impedance (or can be bypassed by a capacitor to make it act like a low impedance); therefore, the end of R_C connected to the battery is at signal ground potential. Since current flows through the resistor in the direction shown by the arrows, the end of R_C connected to the collector will be negative with respect to the end connected to the battery. The polarity of the end of the resistor connected to the collector can be represented by the straight line drawn below the zero axis as shown by curve 1 of Fig. 8A.

Now let us consider what happens when an input signal is applied to the input, as shown in curve 2 of Fig. 8A. At point 1, the input voltage is zero, and the only current flowing through the transistor is caused by the battery voltages. The signal voltage across R_C at that instant is zero and identified at point 1 on curve 3 of Fig. 8A. Now when the input signal swings in a positive direction so that the end of R1 connected to the base is positive and the end connected to ground is negative, the voltage across resistor R1 will be in series with the voltage of battery B1. This will increase the forward bias applied across the emitter-base junction of the transistor. This will increase the number of electrons crossing the emitter-base junction and hence increase the number of electrons flowing through the transistor to the collector. Therefore the number of electrons flowing through resistor R_C will increase, and the voltage drop across the resistor will increase, making the end of the resistor connected to the collector more negative with respect to

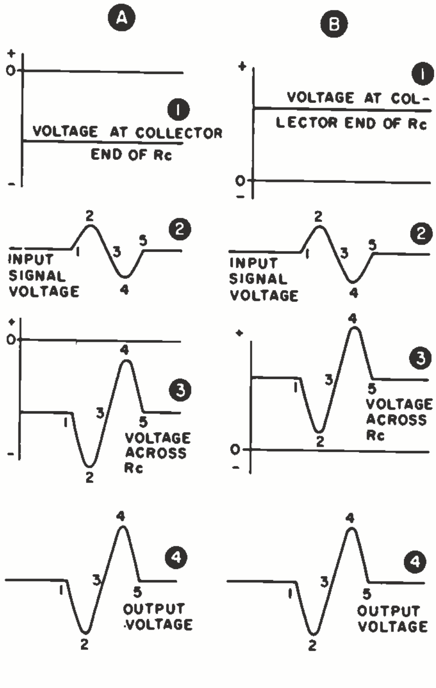


Fig. 8. Voltage waveforms for the circuits of Fig. 7. (A) is for the NPN transistor circuit, and (B) is for the PNP transistor circuit.

the end connected to the battery. Thus the output voltage will swing in a negative direction to point 2 on the output voltage curve, 3 of Fig. 8A.

As the input voltage drops back to zero, the forward bias applied across the emitter-base junction will decrease until, when the signal voltage reaches point 3, the forward bias will be made up of only the battery voltage and zero-signal current will be flowing through R_C. The voltage across R_C will drop to point 3 on the output voltage curve, 3 of Fig. 8A.

When the polarity of the input signal reverses, making the end of R1 that is connected to the base negative and the end connected to ground positive, the input voltage across resistor R1 will oppose the bias voltage applied between the emitter and base by battery B1. Thus the forward bias applied across the emitter-base junction will decrease, and the current flowing across this junction will decrease. When the current flowing across this junction decreases, the collector current, and hence the current flowing through load resistor R_C, will decrease. When the current flowing through this resistor decreases, the voltage drop across it will decrease, and the voltage at the end of the resistor connected to the collector will approach zero. This is represented by point 4 on curve 3 of Fig. 8A.

Finally, when the input signal voltage again drops to zero as shown at point 5 on curve 2 of Fig. 8A, the voltage across R_C will again increase in a negative direction until it reaches point 5 on curve 3 of Fig. 8A.

In curve 4 we have shown the output signal voltage that can be obtained by removing the dc component of the total voltage appearing across the resistor R_C. Again, this dc component can easily be removed by connecting a capacitor in series with one of the output leads.

Compare the output voltage curve shown in 4 of Fig. 8A with the input voltage shown in 2 of Fig. 8A. Notice that when the input voltage swings positive to point 2, the output voltage swings negative to point 2. Similarly, when the input voltage swings in a negative direction to point 4, the output voltage swings in a positive direction to point 4. This means that when the input is going in a positive direction, the output is going in a negative direction. In other words, the output signal voltage appearing across the load resistor R_C is 180° out-of-phase with the input signal voltage appearing across R1. From this we can conclude that when an NPN transistor is used in a common-emitter circuit, the amplified output voltage will be 180° out-of-phase with the input voltage. Also notice when comparing curve 4 with curve 2 that the amplitude of curve 4 is greater than the amplitude of curve 2. In other words, there is a voltage gain in this circuit.

A COMMON-EMITTER PNP AMPLIFIER

Although we have a somewhat different situation in a common-emitter amplifier using a PNP transistor, the net result is the same.

In a common-emitter amplifier using a PNP transistor, as shown in Fig. 7B, the electron current through resistor R_C is in the opposite direction to what it was in the amplifier using the NPN transistor. In the PNP amplifier, the electrons flow from the negative terminal of battery B2, through load resistor R_C into the collector to fill the holes crossing the base-collector junction. The electrons flowing through load resistor R_C set up a voltage drop across it having a polarity such that the end of the resistor connected to the collector is positive with respect to the end connected to the battery, as shown in Fig. 7B. Thus, if we plot the voltage at the collector end of this resistor with respect to the other end we have a curve like the one shown in 1 of Fig. 8B. Here the zero-signal voltage across R_C is represented by a straight line drawn above the zero voltage axis to indicate the fact that this voltage is positive.

Now, when an input signal like the one shown in curve 2 of Fig. 8B is applied across R1, the collector current will vary as before. When the input signal swings positive so that the end of R1 that is connected to the base is positive and the end connected to ground is negative, the voltage across R1 will oppose the voltage of battery B1, thus reducing the forward bias applied across the emitter-base junction. When the forward bias is reduced, the number of holes crossing this junction and traveling through the base, across the base-collector junction to the collector terminal is reduced. If the number of holes reaching the collector is reduced, the number of electrons flowing through R_C to fill the holes reaching the collector will be reduced. Hence the voltage drop across R_C will decrease, in other words drop toward zero. This is shown by curve 3 of Fig. 8B. When the input voltage shown in curve 2 swings in the positive direction, from point 1 to point 2, the output voltage across R_C moves from point 1 to point 2 on curve 3 of Fig. 8B.

When the input signal applied across R1 swings in a negative direction so that the end of R1 that is connected to the base is negative and the end connected to ground is positive, the voltage across resistor R1 will be in series with the voltage of battery B1 and will add to the forward bias applied across the emitter-base junction by the battery. This increased forward bias will result in an increase in the number of holes crossing the emitter-base junction which will, in turn, mean that there will be an increase in the number of holes reaching the collector. If more holes reach the collector, more electrons will have to flow through the output load resistor, R_C, to fill these holes. The result will be that the voltage drop across this resistor will increase above the zero-signal voltage level, and the end of R_C connected to the collector will become more positive with respect to the end connected to B2. This increase in voltage is shown as point 4 on curve 3 in Fig. 8B.

The complete cycle across R_C obtained with an input voltage such as shown in curve 2 is shown in curve 3 of Fig. 8B. When the dc component across resistor R_C is removed, we have the results shown by curve 4 of Fig. 8B. Notice that once again the output signal is 180° out-of-phase with the input signal.

By comparing the NPN amplifier with the PNP amplifier, you can immediately see that, although the action is somewhat different in the two circuits, the net results are the same. In both cases we have voltage amplification, and in both cases we find that in the common-emitter circuit the output signal is 180° out-of-phase with the input. Another way of saying this is that a common-emitter amplifier reverses the phase of the signal.

CHARACTERISTICS OF A COMMON-EMITTER CIRCUIT

The common-emitter circuit is the most important of all transistor circuits; it is by far the most widely used.

It is quite easy to get relatively high gain using the common-emitter circuit. Voltage gains from 80 to 100 are quite easily obtained. You can also obtain a current gain with this circuit.

Although the input impedance of this circuit is not as high as the input impedance of a vacuum tube circuit, it is substantially higher than the input impedance of the common-base circuit. Therefore it can be driven by a much higher impedance device than a common-base circuit. In a typical common-emitter circuit we usually have an input impedance of somewhere between 1000 and 2000 ohms. The output impedance is not quite as high as the output impedance of the common-base circuit, but it is still high, usually having a value of around 20,000 ohms.

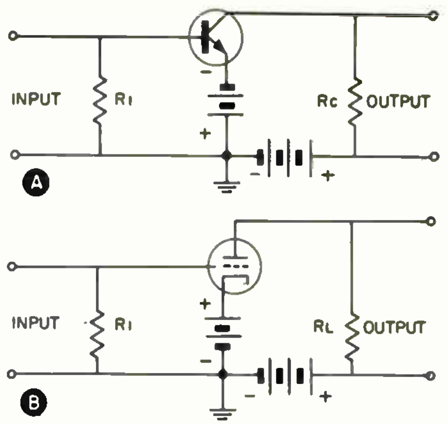


Fig. 9. Comparison of common-emitter transistor circuit (A) and grounded-cathode vacuum-tube circuit (B).

COMMON-EMITTER AND GROUNDED-CATHODE CIRCUITS COMPARED

Because of circuit similarities and performance similarities, the common-emitter circuit can be compared to the grounded-cathode vacuum tube circuit. The grounded-cathode is the most common of all vacuum tube circuits. With it, as with the common-emitter transistor circuit, it is possible to obtain voltage gains of 80 to 100, and the phase is shifted 180°.

In order to help you see the similarity between these two circuits, the common-emitter circuit and the grounded-cathode vacuum tube circuit are shown in Fig. 9. The battery shown in the cathode circuit of the grounded-cathode amplifier is not found in actual practice because a resistor can be used in this circuit to avoid the necessity of this extra battery. You will see later that similar arrangements are used in transistor circuits to avoid the necessity of using two batteries to operate a single stage.

SUMMARY

The common-emitter circuit is the most frequently used transistor circuit. The voltage gain is from 80 to 100 and there is also considerable current gain. It has a medium input resistance and an output resistance of about 20,000 ohms. The output signal is 180° out-of-phase with the input signal.

SELF-TEST QUESTIONS

(g) Is it possible to get a current gain using the common-emitter circuit?

(h) What is the relationship between the input signal voltage and the amplified output signal voltage?

(i) How does the input impedance of the common-emitter circuit compare with the input impedance of the common-base circuit?

(j) Draw schematic diagrams of common-emitter circuits using NPN and PNP transistors. You should do this from memory since it is important that you remember these circuit configurations.

(k) To what vacuum tube circuit can the common-emitter circuit be compared?

The Common-Collector Circuit

The third possible circuit configuration using a triode transistor is the common-collector circuit. In this type of circuit the collector is operated at signal ground potential. Although this circuit is not found as often as the common-emitter circuit, it does have some characteristics that are useful in some special applications.

A common-collector circuit using an NPN transistor is shown in Fig. 10A, and one using a PNP transistor is shown in Fig. 10B. The arrows on the diagram in Fig. 10A show the direction of electron flow through the circuit. The solid arrows in 10B show the electron flow, and the outlined arrow indicates the direction of hole movement through the PNP transistor. Compare the circuits shown in Fig. 10 with the common-base circuits shown in Fig. 1 and the common-emitter circuits shown in Fig. 7. Let's see how each of these circuits works.

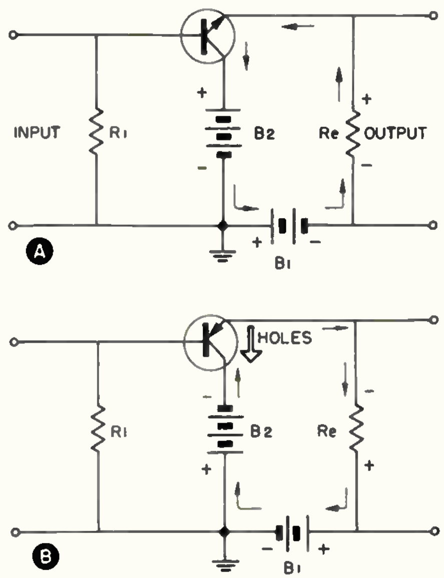


Fig. 10. Common-collector circuits. (A) for an NPN transistor, and (B) for a PNP transistor.

A COMMON-COLLECTOR NPN AMPLIFIER

In the circuit shown in Fig. 10A, electrons flow from the negative terminal of B1 through the emitter resistor Re to the emitter. They flow across the emitter-base junction, through the base, across the base-collector junction, to the positive terminal of B2. Battery B1 biases the emitter-base junction in a forward direction, whereas B2 biases the base-collector junction in a reverse direction.

Electrons flowing through Re set up a voltage drop across Re with the polarity indicated on the diagram. The end of the resistor connected to the emitter is positive with respect to the end connected to the battery. If we plot the voltage at the emitter end of the resistor, it will be like curve 1 of Fig. 11A. The voltage is represented by a straight line drawn above the zero voltage axis to indicate that it is positive with respect to the other end of the resistor.

Now let us consider what happens when an input signal like that shown by curve 2 of Fig. 11A is applied across the input terminals. When this signal swings in a positive direction from point 1 to point 2, the end of R1 connected to the base will be positive and the other end will be negative.

This means that the voltage across R1 will be in series with battery B1 and will add to the forward bias ap-

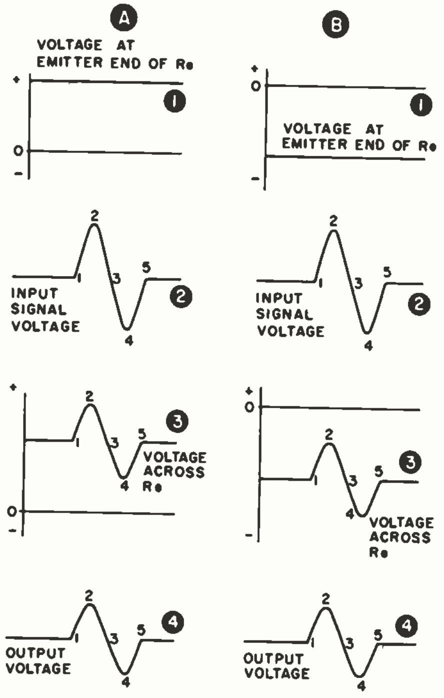


Fig. 11. Voltage waveforms for the circuits of Fig. 10. (A) is for the NPN transistor circuit; (B) is for the PNP transistor circuit.

plied across the emitter-base junction. This will increase the flow of electrons across this junction, and hence increase the current flow through the circuit and through Re. The increase in current through Re will result in an increase in the voltage drop across Re. This means that the end of the resistor connected to the emitter will become more positive with respect to the other end. The change in the voltage is indicated by points 1 and 2 on curve 3 of Fig. 11A.

Let's stop for a minute and consider what happens to the voltage applied across the emitter-base junction at this time. When the volttage across Re increases, it subtracts from the total voltage across the emitter-base junction. You will notice that the voltage across Re opposes the voltage of battery B1. Therefore an increase in the voltage across Re results in an increase in the opposition to the voltage of battery B1. When the input signal swings in a positive direction, the voltage across Re increases, opposing the increase in emitter-base voltage producing it. In other words, we have 100% voltage feedback. All of the amplified voltage appearing across Re opposes the input voltage producing it.

When the voltage across R1 swings in a negative direction as shown between points 3 and 4 on the input voltage curve, 2 of Fig. 11A, the input voltage will oppose the voltage of battery B1. This will reduce the emitter-base bias, resulting in fewer electrons crossing the emitter-base junction. This means that the current flowing in the circuit will decrease, and therefore the voltage across Re will decrease. This is shown between points 3 and 4 on curve 3 of Fig. 11A. This decrease in the voltage across Re will result in a reduction of the opposition of this voltage to the voltage of battery B1. In other words, we again have a situation where the output signal voltage being produced across Re is opposing the input signal voltage producing it.

In curve 4 of Fig. 11A we have shown the output voltage with the dc component removed. Notice that this voltage is in phase with the input voltage. Also notice that this voltage is smaller than the input voltage. Since we have 100% voltage feedback in this circuit, the output voltage will always be less than the input voltage. This situation is similar to the situation in the grounded-plate or cathode-follower vacuum tube amplifier. This transistor circuit is often compared to the grounded-plate amplifier, and we will soon see the similarity between these two circuits.

A COMMON-COLLECTOR PNP AMPLIFIER

In the common-collector circuit using a PNP transistor, the polarity of the voltage across Re is the opposite of what it was with the NPN transistor. Thus the voltage at the emitter end of resistor Re is represented by a straight line drawn below the zero signal axis as shown by curve 1 of Fig. 11B. In curve 2 of Fig. 11B we have shown an input signal similar to the one shown in curve 2 of Fig. 11A. When this input signal swings in a positive direction so that the end of R1 connected to the base is positive and the grounded end is negative, the voltage across R1 will oppose the voltage of battery B1. This will reduce the forward bias applied across the emitter-base junction and reduce the number of holes crossing this junction. You will remember that holes are formed in the emitter by pulling the electrons off the emitter. If fewer holes are formed, fewer electrons will be pulled off the emitter, and hence the current flowing through Re will decrease. When the current flowing through Re decreases, the voltage drop across Re decreases. This can be seen between points 1 and 2 on curve 3 of Fig. 11B.

When the input voltage applied across R1 swings in a negative direction as between points 3 and 4 of curve 2, the end of resistor R1 connected to the base will be negative and the grounded end positive. The voltage across R1 will be in series with battery B1 and will add to the emitter-base forward bias. This will result in an increase in the number of holes crossing the emitter-base junction. More electrons will therefore be pulled out of the emitter to produce additional holes across this junction. The current flowing through Re will increase, resulting in an increase in the voltage drop across Re. This increase is shown between points 3 and 4 on curve 3 of Fig. 11B.

Once again we have shown the output signal voltage in curve 4. Notice this is identical to the output signal voltage obtained with the NPN transistor. Notice that although the basic operation of the two circuits is somewhat different, the net result is the same. In both cases we have 100% voltage feedback, so the output is less than the input. Also notice that in both cases the output signal voltage is in phase with the input signal voltage.

CHARACTERISTICS OF COMMON-COLLECTOR CIRCUITS

The common-collector circuit has several interesting characteristics. It has the highest input impedance of the three circuits. You can see why this is true if you refer to Fig. 10. The input impedance will be the ratio of the input voltage over the input current. The voltage applied across resistor R1 will cause a certain current to flow through it. In addition, this input voltage will cause a signal current to flow from the emitter, through the transistor, to the collector. In the circuit shown in Fig. 10A, part of the electrons travelling from the emitter to the collector will be attracted by the base and hence will flow through R1. However, since the output signal voltage subtracts from the input signal voltage insofar as signal voltage applied between base and emitter is concerned, the actual signal current flowing through the transistor will be quite small. Therefore the total input signal current will be small and this, in turn, will result in the input impedance being high. On the other hand, the output impedance of the transistor will be low. This is due to the fact that the emitter signal current flows through the resistor Re and that very little voltage will be developed across this resistor. As a matter of fact, the voltage cannot be equal to the input signal voltage because if it was it would cancel the signal voltage entirely. Therefore, since the output voltage is small, the ratio of the voltage divided by the current will be small and the output impedance will be low.

In a common-collector circuit the voltage gain is always less than one. This means that the output voltage will always be less than the input voltage. We have already pointed out that this must be true because otherwise the output voltage would completely cancel the input voltage. Of course, this is impossible because it is the input voltage that causes the current change through the transistor which in turn develops the output voltage. All of the output voltage is fed back into the input circuit and therefore we say that we have 100% voltage feedback.

The common-collector circuit has the best stability of the three transistor circuits. This is what you might expect because, with the low output voltage, there is very little voltage to produce feedback into the input circuit which could cause instability or oscillation. Also, there is no phase reversal in the circuit and therefore the output voltage is in phase with the input voltage.

We have redrawn the common-collector circuit in Fig. 12A to make it easier to compare it with a cathode-follower or grounded-plate vacuum tube circuit shown in Fig. 12B. Notice the similarity between the cathode - follower and common collector circuit.

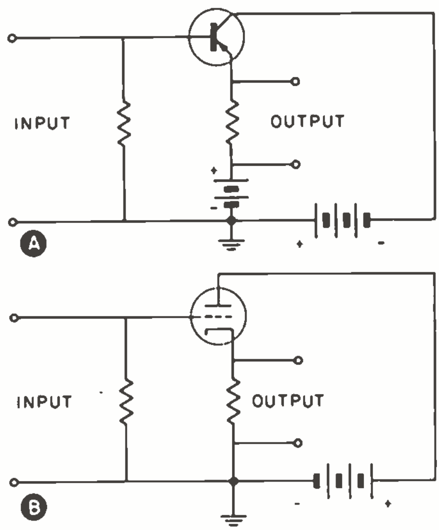


Fig. 12. Comparison of a common-collector transistor circuit (A) and a cathode-follower vacuum-tube circuit (B).

Because the common-collector circuit has a high input impedance and a low output impedance it is often used as an impedance matching device to match a relatively high impedance to a low impedance. An excellent use of this type of circuit is in the video amplifier of color TV receivers. In a color TV receiver a delay line, that slows up or delays the video signal, is used in the video amplifier. This delay line is used so that the brightness signals being fed to the color picture tube will be slowed down a little so that they will arrive at the picture tube at the same time as the color signal. A delay line is a comparatively low impedance device and the common-collector circuit provides an excellent method of matching from the higher impedance video amplifier stages to the low impedance of the delaying line. It is quite likely that when more transistorized color TV receivers are manufactured, this circuit will be very useful in this application.

SUMMARY

It is important for you to understand the three basic transistor circuits. You will find that transistors in commercial equipment will be arranged in one of these three basic circuits.

The most commonly used of the three circuits is the common-emitter circuit. In this circuit, the emitter is common to both the input and the output circuits. It is operated at signal ground potential. In the common-emitter circuit a voltage gain of from 80 to 100 can easily be obtained. In addition, there will be considerable current gain in this circuit. Other important characteristics of this circuit are a medium input resistance, usually somewhere between 1000 and 2000 ohms, and an output resistance in the neighborhood of 20,000 ohms. You should also remember that this is the voltage amplifier circuit that produces a 180° phase shift. In other words, the output signal voltage will be 180° out-of-phase with the input signal voltage.

The common-base circuit is the circuit in which the base is common to both the input and output circuits. It has a very low input resistance but has the highest output resistance of the three basic circuits. The current gain is always less than 1, but this type of circuit is quite stable. In fact, temperature changes have little effect on the operation of the circuit, whereas this is not always true of the common-emitter circuit. There is no phase reversal in a voltage amplifier used in this type of circuit; in other words, the output voltage will be in phase with the input voltage.

The common-collector amplifier is an amplifier in which the collector circuit is common to both the input and output circuits. This circuit has the highest input resistance, but the output resistance is very low; it may be as low as 100 ohms. This is the only one of the three circuits that has a lower output resistance than input resistance. In this circuit the voltage gain is always less than 1 because there is 100% voltage feedback. The stability of this circuit is excellent--the best of the three circuits. Again, there is no phase reversal when this circuit is used; the output voltage is in phase with the input voltage.

SELF-TEST QUESTIONS

(l) What is the phase relationship between the output voltage and the input voltage in a common-collector circuit?

(m) What will the voltage gain of the common-collector circuit be?

(n) What are the relative input and output impedances of the common-collector circuit?

(o) To what vacuum tube circuit can we compare the common-collector circuit?

(p) In which transistor circuits will you find an output voltage that is in phase with the input voltage?

(q) Which transistor circuits give you an output signal voltage that is 180° out-of-phase with the input signal voltage?

(r) Which transistor circuits will give you a voltage gain?

(s) Which transistor circuits will give you a current gain?

(t) Which transistor circuits have a relatively high input impedance?

(u) Which transistor circuit has a low output impedance?

Transistor Characteristics

You will remember from your study of tubes that they have certain important characteristics that tell the technician a great deal about how the tube should perform. The important tube characteristics are the mutual conductance, amplification factor and plate resistance. Likewise, there are certain transistor characteristics that are important to the technician. They enable him to compare one transistor with another and to get an idea of what to expect from a transistor in a certain circuit. All of this information is helpful in determining whether or not a transistor is performing the way it should.

In addition, there are a large number of symbols used in describing transistor performance. Many of these symbols are of interest only to circuit designers and engineers, but the technician should be familiar with the more important ones and be able to evaluate from a transistor manual the important transistor characteristics. In this section of this lesson we are going to cover some of the more important symbols and transistor characteristics.

TRANSISTOR SYMBOLS

As you might expect, the letter I is used to represent current in transistor circuits. When the capital letter I is used, it indicates dc current or rms current. When the small letter i is used, it indicates instantaneous current.

Currents flowing in the various transistor electrodes are identified by means of a letter representing the electrodes. For example, the emitter current is represented by the letter E or e. Base current is represented by B or b and collector current is represented by the letter C or c. Using these symbols, the dc emitter current is designated by the symbol I_E. The rms emitter current is represented by the symbol I_e, and the instantaneous emitter current is represented by the symbol i_e. Similarly, dc base current is represented by I_B, rms base current is represented by I_b, and instantaneous base current is represented by i_b. Collector dc current is represented by I_C, rms collector current is represented by I_c, and instantaneous collector current is represented by i_c.

Two characteristics that are often referred to in transistors are the forward current and the reverse current. The symbol used to represent the dc forward current is I_F, and i_F is used to represent the instantaneous-forward current. The dc reverse current is represented by I_R and the instantaneous-reverse current is represented by i_R.

You will remember that in normal operation a transistor is operated with a forward bias across the emitter-base junction and a reverse bias across the base-collector junction. Thus in an NPN transistor, current can flow from the emitter, across the emitter-base junction, through the base, across the base-collector junction and through the collector to the positive terminal of the battery, placing the reverse bias across the base-collector junction. However, we also point out that there would be at all times some minority carriers crossing the various junctions in a reverse direction. Thus there will be a current flow across the base-collector junction due to holes travelling from the collector across the junction into the base. This reverse current is kept as small as possible because it contributes nothing to the usefulness of the transistor. As a matter of fact, the current crossing the junction tends to heat the junction and cause a number of undesirable effects. Transistor manuals often list the reverse current across the collector-base junction. The current that is listed is the current that will flow across the junction when the junction is reverse biased and the emitter is open circuited. This dc current is represented by the symbol - I_{CBO}. The letters CB indicate that the current is across the collector-base junction in the reverse direction. The letter O indicates that the other electrode, the emitter, is open. This symbol is so widely used, that it is often abbreviated I_{CO}.

Groups of symbols are used in this manner to indicate other transistor current. For example, the symbol I_{CEO} is used to represent the dc collector current with the collector junction reverse biased and the base open circuited.

There are other symbols used in conjunction with transistors, but the ones covered in this section are the most important ones for the technician to remember, along with the few new ones we will cover in the next section. If you read through this section carefully and understand how these symbols are put together, the chances are that you will be able to figure out any that you are likely to encounter that will be of importance to you.

CURRENT GAIN

As we mentioned, a transistor is primarily a current operated device. Its ability to amplify is due to the fact that it can transfer a current from a comparatively low resistant circuit to a higher-resistant circuit. One of the important characteristics of a transistor is its current gain.

Since the current gain that can be obtained with a transistor depends upon the circuit in which the transistor is used, two symbols are used for current gain. These symbols are the Greek letters alpha (α) which represents the current gain in a common-base circuit, and the Greek letter Beta (β), which is used to represent the current gain in a common-emitter circuit. The two are interrelated; let's see how and exactly what each symbol means.

Alpha.

Alpha is equal to the change in collector current divided by the change in emitter current needed to produce this change in collector current. This is often represented by the symbols:

$$\alpha = \frac{\Delta\ Ic}{\Delta\ Ie}$$

The small triangles are Greek letter deltas, which are used to indicate a change; in this case, a change in current.

You will remember that in a common-base circuit the current gain is less than 1 because the change in collector current is slightly less than the change in emitter current. This is due to the fact that not all of the carriers crossing the emitter-base junction reach the collector. Some of them are attracted to the battery in the emitter-base circuit, and some of them are lost through recombination in the base. Thus the number of carriers reaching the collector will be slightly less than the number of carriers crossing the emitter-base junction. However, in a good transistor the majority of the carriers do reach the collector so that the current gain in the common-base circuit is close to 1. Typical values run around .95, which indicates that 95% of the carriers crossing the emitter-base junction reach the collector.

Transistor manufacturers often list the alpha of a transistor in the transistor characteristics. This will immediately tell you what current gain can be obtained from the transistor used in a common-base circuit. Also, as you will see later, you can determine from this figure the current gain that will be obtained in the same transistor in a common-emitter circuit. Another characteristic often given is the alpha cut-off frequency. This is the frequency at which the current gain of the transistor in the common-base circuit drops to 70.7% of what it is at lower frequencies.

Beta.

You will remember that in the common-emitter circuit we had a current gain. This means that the value of beta will always be greater than 1.

Beta is defined as the change in collector current divided by the change in base current. It is often represented by the expression:

$$\beta = \frac{\Delta \, Ic}{\Delta \, Ib}$$

Again, the small triangles are used to indicate a change in current. Typical values of beta may run as high as 80 or 100.

Beta is another characteristic frequently found in transistor specifications. If you know the beta of a transistor, you immediately know the current gain that the transistor will give when it is used in a common-emitter circuit. From this, as you will soon see, you can also determine the current gain that will be obtained in the common-base circuit, if the value is not given in the characteristics. The beta cut-off frequency is the frequency at which the current gain of the transistor in a common-emitter circuit drops to 70.7% of what it is at lower frequencies.

Converting Values.

Manufacturers often give either the alpha or the beta of a transistor, but seldom both. Sometimes when the alpha is given you want to know the value of beta and vice versa. Actually it is quite easy to convert from one to the other. If you know the alpha of a transistor you can find beta from the formula:

$$\beta = \frac{\alpha}{1 - \alpha}$$

If you know the beta of a transistor you can find alpha from the formula:

$$\alpha = \frac{\beta}{1 + \beta}$$

Now let's work a couple of examples to see how easy it is to convert from one value to the other. Let's assume that we have a certain transistor and the manufacturer lists the value of alpha as .95. Let's find the value of beta.

Starting with the formula:

$$\beta = \frac{\alpha}{1 - \alpha}$$

We substitute .95 for alpha and get:

$$\beta = \frac{.95}{1 - .95} = \frac{.95}{.05}$$

We can eliminate the decimals in this division by moving both the decimal points two places to the right so we have:

$$\beta = \frac{95}{05} = \frac{95}{5} = 19$$

Thus if alpha is equal to .95, beta will be equal to 19.

Now let's assume we have been given the value of beta as 19 and see what value of alpha we get. We know it should be .95. Starting with the formula:

$$\alpha = \frac{\beta}{1 + \beta}$$

we substitute 19 for beta and get:

$$\alpha = \frac{19}{1 + 19} = \frac{19}{20}$$

to get the value of alpha we need only divide 19 by 20:

```
     .95
20/19.0
   18 0
    1 00
    1 00
```

so we get a value of .95 for alpha if we start with 19 for beta.

Sometimes the gain of a transistor is referred to as a forward current-transfer ratio. In other words, instead of referring to the gain in a common-base circuit we say that the forward current-transfer ratio is alpha; we know this will always be less than one. Similarly, in the common-emitter circuit we refer to beta as the forward current-transfer ratio which, in effect, is the same thing as the current gain of the transistor.

Another characteristic or term that is frequently used in transistor manuals is the gain-bandwidth product. The gain-bandwidth product is the frequency at which beta equals one. In other words, as the frequency at which the transistor is used is increased, the current gain in the common-emitter circuit will drop off. At some frequency beta will be equal to one and this is called the gain-bandwidth product.

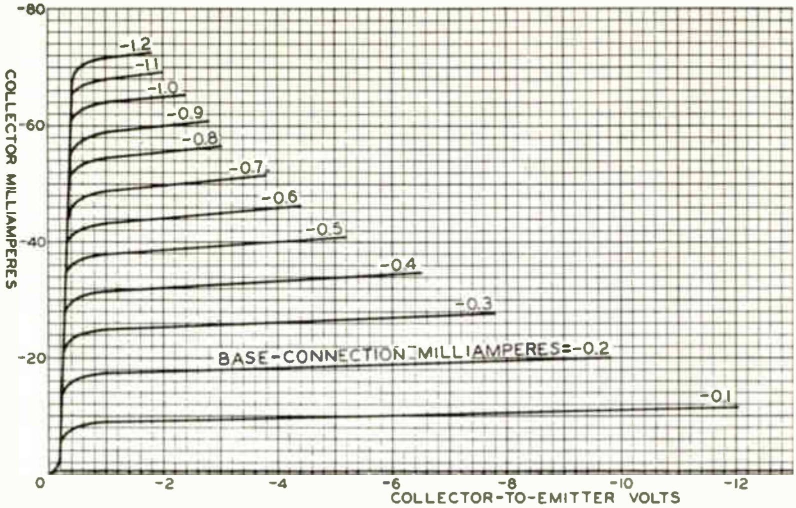
CHARACTERISTIC CURVES

When you studied vacuum tubes, you learned that characteristics are used to supply information about the manner in which a given tube performs. By means of curves a great deal of information about the tube can be condensed and presented in a convenient form. It is possible to see what the tube will do with different operating voltages and under different operating conditions.

Transistor characteristic curves are used for exactly the same reason. The curves give information about the way in which a transistor will perform in a convenient compact form.

A typical set of characteristic transistor curves is shown in Fig. 13. These curves are for a transistor used in a common-emitter circuit. They indicate what the collector current will be with different values of collector-to-emitter voltages and with different base currents. This type of curve is the most widely used transistor characteristic curve.

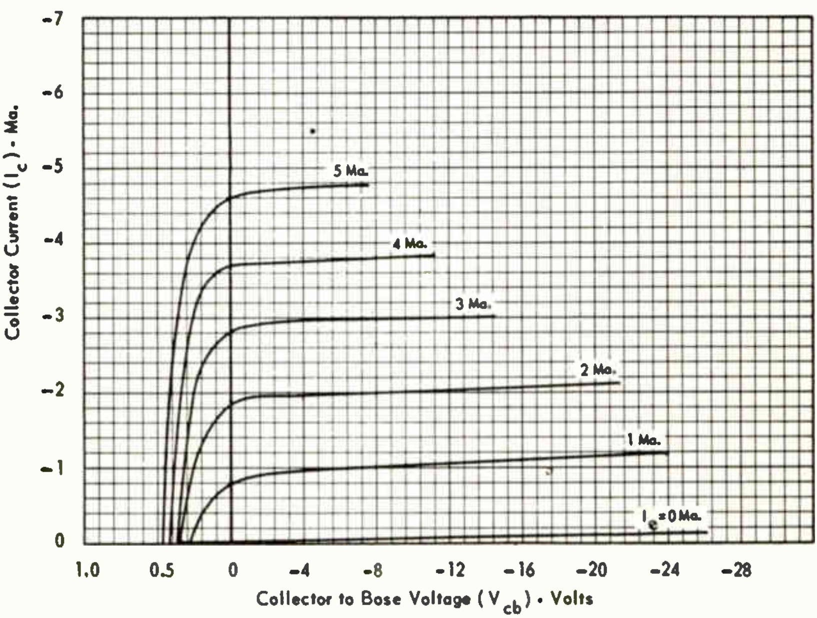
The curves shown in Fig. 13 can be used to determine the beta of a transistor for different collector-to-emitter voltages. For example, at a voltage of -4 volts, notice that when



Courtesy RCA

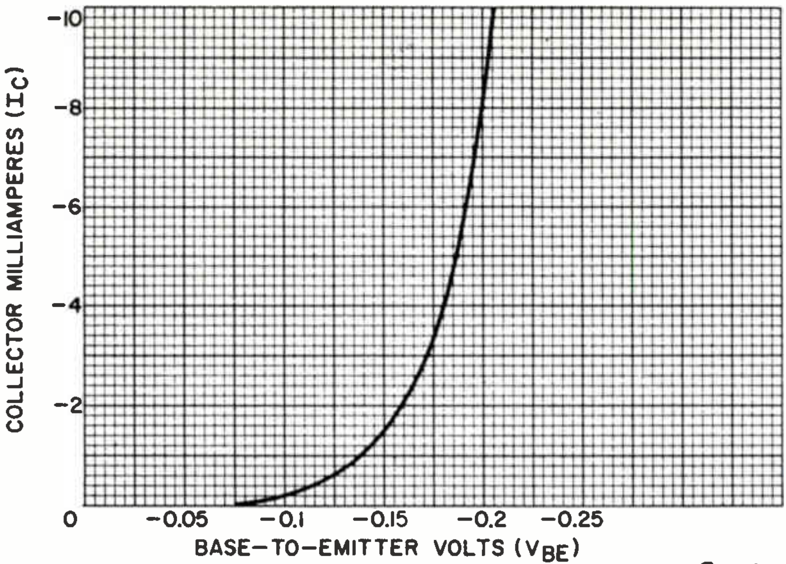
Fig. 13. Typical collector characteristic curves for a transistor used in a common-emitter circuit for different base currents.

the base current changes from -.3 milliamperes to -.4 milliamperes the collector current changes from approximately -26 milliamperes to -33 milliamperes. Thus for the change in base current of -.1 ma we have a change in collector current of -7 ma. Therefore:



Courtesy Raytheon Mfg. Co.

Fig. 14. Typical characteristic curves for a PNP transistor used in a common-base circuit for different emitter currents.



Courtesy RCA

Fig. 15. A typical transfer curve.

$$\beta = \frac{-7}{-.1} = \frac{70}{1} = 70$$

Another set of transistor characteristic curves for a PNP transistor is shown in Fig. 14. This set of curves is for a transistor used in a common-base circuit. Curves of this type are not given as often as those shown in Fig. 13 for the common-emitter circuit, since the common-emitter circuit is more widely used than the common-base circuit.

However, manufacturers often release detailed specification sheets on certain transistor types where curves of this type are given.

Another characteristic curve that is frequently given is the transfer characteristic curve, shown in Fig. 15, which enables you to find the collector current for different base-to-emitter voltages at a fixed collector-to-emitter voltage. An indication of the collector current change that might be expected with a given input voltage can be obtained from a curve of this type. The curve also indicates the range over which the transistor is comparatively linear. For example, with the base-to-emitter voltage of 100 millivolts, you can see that the collector current changes will not be linear with changes in input voltage. However, with a fixed base-to-emitter voltage of about -175 millivolts, small changes in this voltage will result in linear current changes.

SUMMARY

In this section of the lesson we have touched briefly on some of the important transistor characteristics. You learned that alpha is a term used to represent the current gain of a transistor in a common-base circuit. The alpha of a transistor is always less than 1, which means that the current gain of a transistor in a common-base circuit is always less than one.

Beta is the symbol used to represent the current gain of a transistor in a common-emitter circuit. The common-emitter circuit always has a current gain, and therefore the value of beta will always be greater than 1. Remember also that the value of beta is not constant; it depends to some extent upon the collector current.

You studied a number of important symbols that are used in transistor manuals to describe transistor performance and in many technical bulletins on transistor circuitry.

The characteristic curves studied in this lesson are typical of characteristic curves issued by transistor manufacturers. They give you an indication of what the collector current will be for different values of collector voltage and different values of emitter or base currents depending upon whether the transistor is used in a common-base or a common-emitter circuit.

Transistors are newcomers to the electronics field compared to vacuum tubes. As a result, from time to time you may see different types of characteristic curves on transistors issued by manufacturers. However, by studying the curves and the information given by the manufacturer on the transistor you can usually learn a great deal about the transistor and the type of performance you would expect from it, and from this information you can evaluate its performance in the circuit to determine whether or not it is operating properly. Being able to do this is a great help to a technician because he is able to determine whether or not he is getting all he should be able to get out of a particular circuit.

SELF-TEST QUESTIONS

(v) What are the symbols used to represent the dc currents in the various transistor electrodes?

(w) What is meant by the symbol I_{CBO}?

(x) Based on the transistor symbols you studied in this section of this lesson, what would you expect the symbol V_{CE} to mean?

(y) What is the Greek letter α used to represent in transistor circuitry?

(z) What is the maximum value that α can have?

(aa) What is meant by the forward current-transfer ratio in a common-base circuit?

(ab) What symbol is used to represent the current gain of a transistor in the common-emitter circuit?

(ac) What is meant by the alpha cutoff frequency?

(ad) What is the gain bandwidth product of a transistor?

(ae) If the beta of a transistor is 24, what is the value of alpha?

(af) Find the beta value of a transistor if alpha is equal to .99.

Typical Transistor Circuits

So far the circuits we have discussed have been basic transistor circuits. The circuits could be used in essentially these forms, but it is more convenient to modify these circuits slightly in actual usage.

The disadvantage of the circuits we have shown is that they use two separate batteries, one to supply the forward bias required across the emitter-base junction and a second battery to supply the reverse bias required across the base-collector junction. In actual practice, these circuits are modified so that both voltages can be obtained from a single voltage source. In battery-operated equipment this eliminates the need for a second battery. In equipment that operates from a power line, it simplifies the power supply somewhat so that the power supply can be arranged to provide either a positive or a negative voltage with respect to ground instead of having to supply both positive and negative voltages.

In this section of this lesson we will look at a number of typical transistor amplifier stages.

AUDIO AMPLIFIERS

A typical common-emitter audio amplifier circuit using an NPN transistor is shown in Fig. 16. Notice that the circuit is somewhat different from the common-emitter circuits you have seen previously inasmuch as only one battery is used in this circuit. Notice that the negative terminal of the battery is connected to the emitter. The positive terminal of the battery is connected to the base through the resistor R1. This would tend to make the base positive with respect to the emitter, which is what we want in order to place a forward bias across the emitter-base junction. You will remember that some of the electrons travelling from the emitter across the junction into the base leave the transistor through the base connection. These electrons will flow through the resistor R1 and set up a voltage drop across the resistor having the polarity shown. Although the number of electrons flowing through R1 will be very small, by using a large value resistor for R1 we can develop considerable voltage drop across it. Therefore even though the base is connected directly to the positive terminal of the battery through R1, the voltage drop across R1 subtracts from the battery voltage so that the net emitter-base voltage is quite small. Usually the base will be positive with respect to the emitter by only a few tenths of a volt, which is all that is required to forward bias the junction.

The collector is connected to the positive terminal of the battery through R2. The value of R2 is chosen

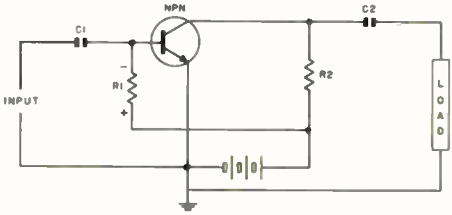


Fig. 16. A common-emitter audio amplifier using an NPN transistor.

so that the current flowing through it will result in a voltage drop across the resistor that is less than the voltage drop across R1. This means that the collector is positive with respect to the base, which is again the condition required to place a reverse bias across the base-collector junction.

In operation, the input signal causes the base-emitter forward bias to vary. This in turn causes the number of electrons crossing the emitter-base junction to vary and hence the collector current varies. The varying collector current flowing through the resistor R2 will result in an amplified signal voltage being developed across R2. This amplified signal voltage is fed to the load through the coupling capacitor C2.

In a typical circuit you will often find that both C1 and C2 are electrolytic capacitors. Electrolytic capacitors are used because the impedances found in transistor circuits are much lower than those found in vacuum tube circuits. You will remember that the capacitor C1 will act as a voltage divider along with R1 and divide the input signal in such a way that part of it will appear across the capacitor and part of it will appear across the transistor input impedance. The amount that will be lost across the capacitor will depend upon its reactance in comparison with the input impedance of the transistor. Therefore, to keep the reactance of the capacitor low, a large value capacitor is usually used.

Fig. 17 is a schematic of a common-emitter amplifier using a PNP transistor. Notice that the circuit is identical except that the battery is reversed. In this circuit, the nega-

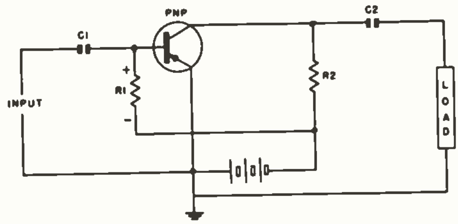


Fig. 17. A common-emitter audio amplifier using a PNP transistor.

tive terminal of the battery is connected to the base of the transistor through R1. This places a negative voltage on the base with respect to the emitter, which will forward bias the emitter-base junction. As in the previous case, however, the entire voltage will not be applied across the emitter-base junction because of the voltage drop across R1. Some electrons will flow from the negative terminal of the battery through R1 and into the base of the transistor to fill holes. This will result in a voltage drop across R1 having the polarity shown in the diagram. This voltage drop will subtract from the battery voltage so that once again the forward emitter-base bias is only a few tenths of a volt.

In this circuit, as in the previous circuit, the voltage drop across R2 will be less than the voltage drop across R1, and therefore the collector will be negative with respect to the base so that we will have a reverse bias across the base-collector junction.

The operation of this circuit is essentially the same as the operation of the circuit shown in Fig. 16. The only exception is that in the transistor holes will be the majority carriers whereas in the preceding NPN transistor, electrons were the majority carriers.

Bias Stabilization.

In the circuit shown in Fig. 16 we have used an NPN transistor. In the collector region, in addition to free electrons which are the majority carriers, there will be some free holes. These holes, which are the minority carriers, will tend to cross the collector-base junction and flow into the base of the transistor. Some of the electrons travelling from the emitter across the emitter-base junction into the base will fill these holes. This is one of the reasons why the collector current will always be less than the emitter current. Some of the holes, however, will be filled by electrons that would normally leave the transistor at the base and flow through R1.

If the temperature of the transistor increases, the resistance of the base-collector junction will decrease, and this will allow additional holes to cross from the collector into the base. This, in turn, will reduce the number of electrons flowing through R1 which will reduce the voltage drop across it.

If the voltage drop across R1 is reduced, the effective emitter-base forward bias will be increased because this bias is equal to the battery voltage less the voltage drop across R1. The increase in forward bias across the emitter-base junction will result in a higher emitter current. This, in turn, will result in a higher current across the emitter-base junction, through the base, and across the base-collector junction. The increase in current across the base-collector junction will result in a further increase in the temperature of this junction which in turn will reduce the resistance again, causing the number of minority carriers crossing the junction to increase still further. This action will continue until eventually the current through the transistor becomes so high that the transistor is destroyed.

The simplest way to overcome this problem is to add a bias-stabilizing resistor in series with the emitter lead. A typical circuit in which this has been added is shown in Fig. 18. The bias-stabilizing resistor is marked R3 on the diagram. Capacitor C3 bypasses the signal around this resistor to prevent degeneration.

If a transistor heats and the base-collector junction resistance is lowered in this circuit, we will have the same situation as before; the resistance will go down and the minority carriers crossing from the collector into the base will increase. This will cause the current through R1 to go down which in turn will reduce the voltage drop across the resistor and tend to increase the emitter-base forward bias. However, the increase in forward bias will tend to cause the emitter current to increase. This increase in current will result in an increase in the voltage drop across R3. Since this voltage subtracts from the forward bias, it will tend to keep the forward bias across the emitter-base junction reasonably stable. Although there will be some increase in current, it is usually not large enough to damage the transistor.

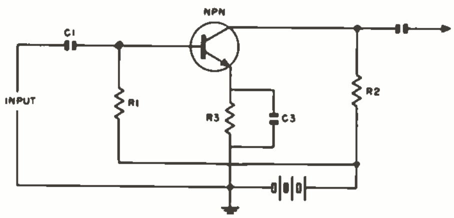


Fig. 18. A common-emitter circuit with a bias stabilizing resistor in the emitter circuit.

Common-Base Amplifier.

The common-base circuit is sometimes found in voltage amplifier circuits. This circuit is particularly useful in TV receivers where some voltage amplification is required without a phase shift. You will remember that with a common-emitter type of amplifier circuit there is a 180° phase shift, whereas in a common-base circuit there is no phase shift.

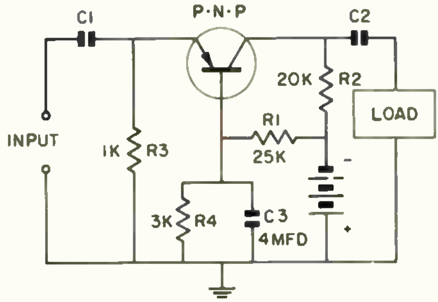


Fig. 19. A single-battery common-base amplifier with bias stabilization.

The circuit shown in Fig. 19 is a common-base circuit modified for use with a single battery. R1 and R4 make up a voltage divider which is connected across the battery. The forward bias, which is required for the emitter-base junction, is developed across R4.

The emitter resistor, R3, serves two purposes in the common-base amplifier. First, it is the impedance across which the input signal is developed. Second, it acts as a bias-stabilizing resistor. As a bias-stabilizing resistor, it works exactly like resistor R3 in Fig. 18, reducing the forward bias of the emitter-base junction if the emitter current tends to increase.

Resistors R1 and R4 are chosen such that the battery current through R1 and R4 is much greater than the base current which flows only through R1. Thus any variations in base current will have little effect on the base voltage which appears across R4. Capacitor C3 bypasses R4 and assures that the base is grounded for ac signals.

RF AMPLIFIERS

Radio frequency amplifiers are almost always tuned amplifiers. They make use of series-resonant and parallel-resonant circuits so that they can be tuned to accept one signal and reject all others. If a low resistance is connected across a parallel-resonant circuit, the resistance will load the circuit so that its selectivity will be destroyed. In other words, it will be unable to select one signal and reject another. This is one of the problems that we face with transistor rf amplifiers. Transistor amplifiers must take into account the fact that transistors are relatively low impedance devices and if they are connected directly across a resonant circuit they will load the circuit and reduce the selectivity.

A typical transistor i-f amplifier is shown in Fig. 20. Although this

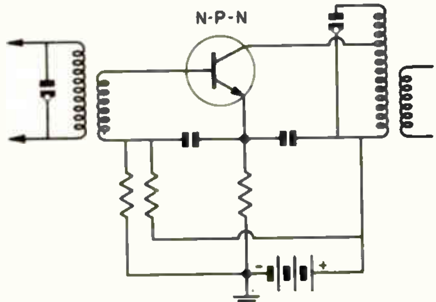


Fig. 20. Transistor input impedance matched by a step-down input i-f transformer and output impedance matched to tuned circuit by tapping down on coil.

is used as the intermediate frequency amplifier in a radio receiver, it is still a radio frequency amplifier since it is designed to amplify radio frequency signals.

Notice that instead of using a double-tuned input transformer, as used in i-f stages using tubes, a single-tuned transformer is used. The secondary of the i-f transformer has far fewer turns than the primary. Therefore the transformer acts like a step-down transformer. The low input impedance of the transistor does not load the primary of the transformer which is a parallel-resonant circuit. Insofar as the primary is concerned, the step-down transformer effectively matches it to the transistor so that excessive loading is avoided.

In the output circuit, the collector is connected to a tap on the primary winding of a second i-f transformer. The lower end of this transformer primary is connected to the positive terminal of the battery and is effectively at signal ground potential. The upper end of the transformer is at some comparatively high impedance with respect to ground. Somewhere along the transformer winding there will be a point at which the impedance of the transformer winding is equal to the output impedance of the transistor. The idea is to connect the collector to this point and by so doing match the collector to the primary winding of the transformer to get maximum signal transfer and at the same time avoid loading the parallel-resonant circuit.

As a technician, you will not have to be concerned about finding the correct impedance point at which to connect a transistor to an i-f transformer; the engineers who designed the set will have taken care of this for you. The important point for you to see is how the transistor is connected into the circuit and to understand why this provision has been made.

One disadvantage of the circuit shown in Fig. 20 is that it is somewhat unstable. This is due to the capacity and resistance across the base-collector junction. Energy can be fed from the collector circuit to the base which is connected directly to the input circuit. If the energy fed back into the input circuit is of the correct phase to reinforce or add to the input signal, the transistor may go into oscillation and generate a signal of its own.

Neutralization.

Two circuits that are used to overcome this undesirable effect are shown in Fig. 21. In both of these circuits neutralization is used to overcome the undesirable effects of feedback.

In the circuit shown in Fig. 21A, tap 2 on the primary of the output i-f transformer is operated at signal ground potential. This tap on the transformer is grounded through the .01-mfd capacitor connected from terminal 2 of the transformer to the emitter. The collector is connected to terminal 1 and signal currents flowing between terminals 1 and 2 will induce a voltage in the portion of the winding between terminals 2 and 3. Thus a voltage is set up at terminal 3 having the opposite polarity to the voltage in terminal 1. This voltage is fed through C1 back to the base of the transistor.

Now, if sufficient energy is fed from the collector of the transistor across the base-collector junction to the base to cause instability, a signal with the opposite polarity is fed through C1 to the base of the tran-

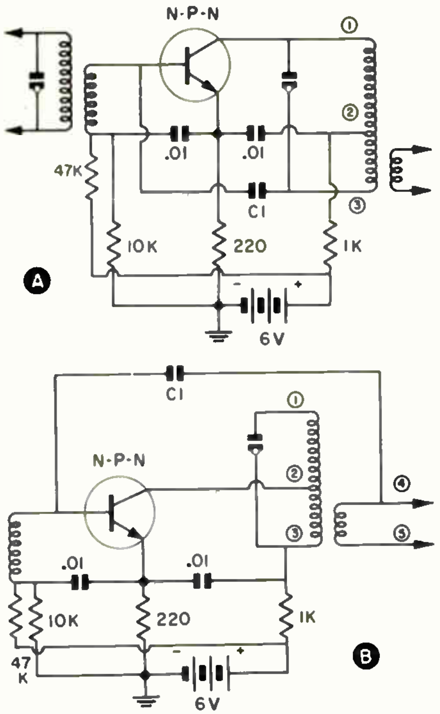


Fig. 21. Two methods of neutralizing a transistor i-f stage.

sistor. These two signal voltages tend to cancel each other so that there is not sufficient net feedback to the base to cause oscillation.

In this circuit the tuned part of the output i-f transformer consists of the entire winding between terminals 1 and 3. You will notice that this winding is tuned by a capacitor so that they form a parallel resonant circuit. Since terminal 2 is operated at ground potential and the collector is connected to terminal 1, the transistor is connected across only a portion of the i-f transformer so that it does not load the transformer excessively and reduce the selectivity of the parallel resonant circuit. At the same time, with this arrangement the instability that was encountered in the circuit shown in Fig. 20 is avoided.

The circuit shown in Fig. 21B is similar to the circuit shown in Fig. 20 except that neutralization is added by means of capacitor C1 connected from the secondary of the output i-f transformer back to the base. The polarity of the signal voltage fed through this capacitor to the base is opposite to the polarity of the signal voltage fed back through the transistor itself across the base-collector junction, so that these two signals will tend to cancel.

When a circuit like the one shown in Fig. 21B is used, the neutralizing capacitor C1 will be somewhat larger than the neutralizing capacitor C1 shown in Fig. 21A because there is quite a step-down in voltage between the primary and secondary windings of the i-f transformer. Since C1 in Fig. 21B is connected to the secondary, the voltage will be considerably lower at this point than it is at the primary. Therefore, a larger capacity is needed in order to feed sufficient voltage back to the base to prevent oscillation. The circuit shown in Fig. 21B is somewhat less critical than the circuit shown in Fig. 21A.

The gain obtained in an i-f stage using a transistor is considerably less than that obtained in an i-f stage using a vacuum tube. As a result, many receivers using transistors in the i-f amplifier have two or more i-f stages. Most modern receivers using vacuum tubes, on the other hand, have only one i-f stage.

OSCILLATOR AND MIXER STAGES

You will remember that modern radio receivers are superheterodyne receivers. In the superheterodyne receiver, the signal being received is mixed in the mixer or converter stage with a signal generated in a local oscillator. Mixing these two signals together results in the formation of two new signals, one with a frequency equal to the sum of the two frequencies and the other with a frequency equal to the difference between the two. Both signals are modulated with the original intelligence being transmitted by the broadcast station.

In modern receivers the difference signal is amplified by the i-f amplifier because the i-f amplifier is tuned to this frequency. This signal is then fed to a detector where the audio signal is separated from the carrier, and then the audio signal is amplified further and fed to a loudspeaker.

Typical mixer and oscillator circuits are shown in Fig. 22. Here an NPN transistor is used in a common-emitter mixer circuit. The primary winding of the loop antenna, L1 on the diagram, is inductively coupled to L2. The signal picked up by the loop is fed to L2 and then applied

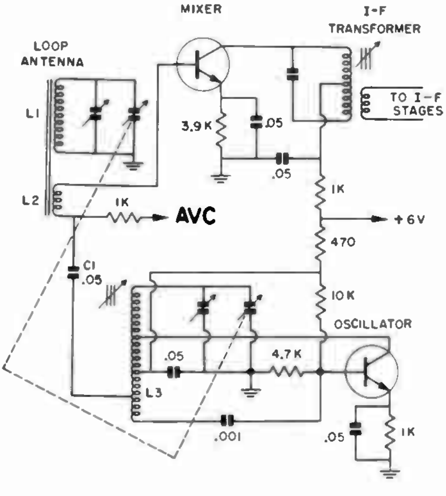


Fig. 22. Oscillator and mixer stages using separate transistors.

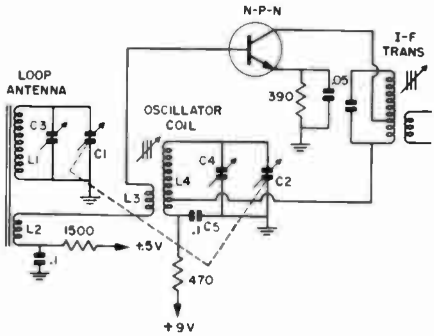


Fig. 23. A converter stage using one transistor as both oscillator and mixer.

between the base and ground. This arrangement prevents the input impedance of the transistor from loading the tuned circuit. The signal from the local oscillator is fed through C1 to L2 and hence to the base of the mixer. In the mixer, the two signals beat together to produce the desired i-f signal.

The oscillator circuit uses another NPN transistor in a common-emitter circuit. This circuit is one form of the Hartley Oscillator which you have seen used previously with vacuum tubes.

The circuit shown in Fig. 22 works well and seldom gives trouble; it's main disadvantage is that two transistors are required. You are not likely to see this arrangement except in some of the more expensive transistor receivers. The usual practice is to use one transistor as both the mixer and the oscillator.

A converter stage, where a single transistor is used as both the mixer and oscillator, is shown in Fig. 23. Here an NPN transistor is used in a common-emitter circuit. In this circuit the signal is picked up by the loop antenna, which is inductively coupled to L2. The signal voltage is induced in series with L2, and this is applied to the base of the transistor through L3, and to ground through the .1-mfd capacitor.

At the same time, energy is fed from the collector into the primary of the i-f transformer and from the primary of the i-f transformer back to L4. L4, C2, C4, and C5 make up the oscillator tank circuit. L4 is inductively coupled to L3, and energy is fed from L3 to the base of the transistor. Thus we have two signals fed to the base of the transistor, the incoming signal picked up by the loop antenna, and the signal generated in the oscillator tank circuit. These two signals are mixed in the transistor and the resulting difference-frequency signal is selected by the i-f transformer and fed to the first i-f amplifier.

Another transistor mixer-oscillator circuit is shown in Fig. 24. This is quite an interesting circuit inasmuch as the transistor is used in both a common-base and a common-emitter circuit at the same time.

In this circuit the incoming signal induced in L3 is applied to the base

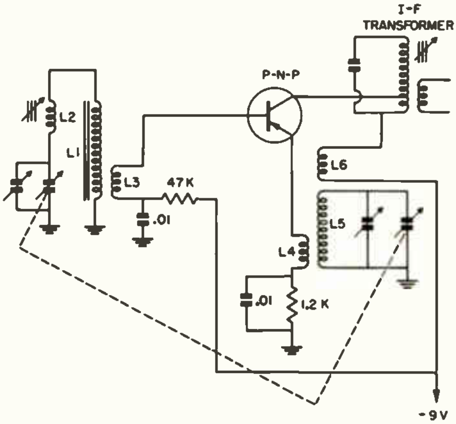


Fig. 24. Converter stage using one transistor connected in a common-base circuit for the oscillator signal, and a common-emitter circuit for the rf signal.

of the transistor. As far as the incoming signal is concerned, the emitter is essentially at ground potential and hence the mixer operates as a common-emitter mixer.

The oscillator signal, on the other hand, is coupled to the feedback winding L4 in the emitter circuit. Therefore, as far as this signal is concerned, the base is the common element and the oscillator operates as a common-base circuit. In spite of the fact that the two signals are fed into different elements of the transistor, mixing still takes place because of the non-linear characteristic of the emitter-base junction.

SUMMARY

The circuits we have studied in these two sections of this lesson are basic transistor circuits. Transistors are found in circuits other than these. Other common circuits in which transistors are found are multivibrators and switching circuits such as used in electronic equipment. However, the circuits that we have studied in this section are what might be considered basic circuits; you should learn these circuit configurations before going on to more complex circuits.

Do not expect to remember what these circuits look like simply by taking a quick look at the diagram in the lesson texts. You should take the time to draw each of these circuits yourself. Draw a circuit two or three times, copying it from the book, then close the book and try to draw it yourself. The chances are that the first time you try to do this you will find that you cannot reproduce the circuit, but after two or three attempts you should be able to do so. Knowing what these basic circuits look like will be a big help to you. You might not understand what all the parts are used for at this time, but we are going to study all of these circuits again in later lessons. If you are familiar with the general circuit configuration, it will be a big help to you because you will know what the circuit looks like and hence will be ready for the next step, that of learning what each and every part in the circuit is used for.

SELF-TEST QUESTIONS

(ag) In the circuit shown in Fig. 16, why is the emitter-base forward bias considerably less than the battery voltage?

(ah) Why is it necessary to provide bias stabilization in transistor amplifiers?

(ai) What advantage does the common-base amplifier circuit offer over the common-emitter circuit?

(aj) In an i-f amplifier such as the one shown in Fig. 20, why is the secondary of the input i-f transformer untuned and shown as having fewer turns than the primary winding?

(ak) What is the purpose of connecting the collector in the transistor circuit shown in Fig. 20 to a tap on the primary winding of the output i-f transformer?

(al) How are the effects of collector-base signal feedback eliminated in transistor amplifiers?

(am) What type of circuits are used in the converter stage shown in Fig. 24?

A Typical Transistor Receiver

One important use for transistors is in portable receivers. Transistors are ideally suited for portable equipment because of their small size and modest current and voltage requirements.

A typical portable receiver is shown in Fig. 25. This receiver uses circuits similar to those you have already studied in this lesson. We will run through this receiver quickly to help you see how the various stages you studied are used together in a complete receiver.

For the purpose of study, we will divide the receiver into two sections, the rf section and the audio section. In the rf section we will include all the stages where the signal present is not an audio signal.

THE RF SECTION

The rf section of this receiver consists of three stages: the converter, which is actually a combination mixer and oscillator, two i-f stages, and the second detector, which is also part of the audio section. The transistors used in these stages are NPN transistors. You will remember that we can identify these transistors as NPN transistors because the arrow used to identify the emitter is pointing up, away from the base.

The converter stage in this receiver is similar to the mixer-oscillator stage shown in Fig. 23. Here we have a transistor used in a common emitter circuit performing the functions of both mixer and oscillator. The primary winding of the i-f transformer T2 is tapped, and the collector is connected to the tap. This is done to avoid loading the high Q i-f transformer and to provide maximum power transfer. Maximum power transfer can be obtained only when impedances are matched. The load impedance of this transistor is quite low, probably about 30,000 ohms, whereas that of the resonant circuit is over 500,000 ohms so the transistor collector is connected to the tap.

The i-f transformer T2 is a stepdown transformer. The input impedance of the first i-f stage is very low, approximately 50 ohms. A stepdown transformer is required in order to match the high-impedance primary circuit to the low-impedance secondary circuit.

The secondary of transformer T2 is tapped. The center tap, marked 5 on the diagram, is at signal ground potential because it is connected to ground through the .1-mfd capacitor C6. This capacitor is in effect a direct connection at the i-f signal frequency. Signal voltages fed from the emitter through the .1-mfd capacitor C5 will cause a current flow through the part of the secondary between terminals 3 and 5. The current flowing through this half of the secondary will set up a field which will induce a voltage between terminals 2 and 5. This voltage is applied between the base and ground and neutralizes the i-f stage. This voltage applied between the base and ground because of the signal fed back through C5 is out-of-phase with the signal voltage fed from the collector through the transistor back to the base, so the two signals cancel.

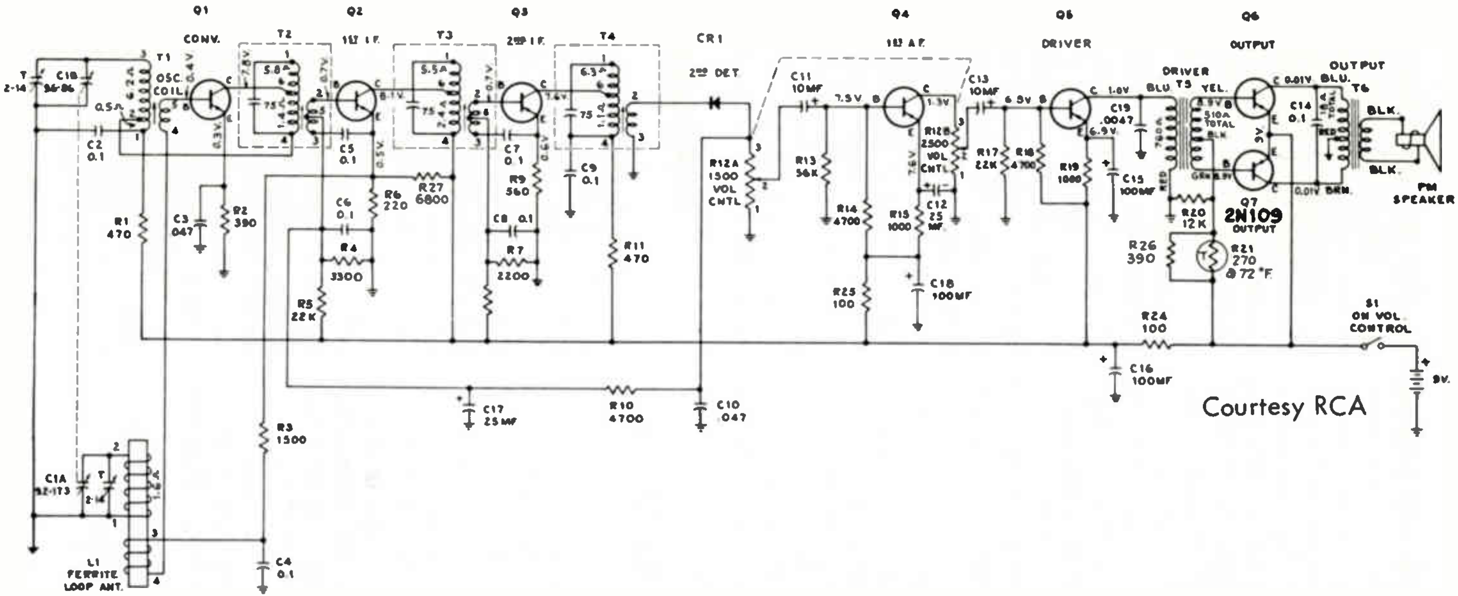


Fig. 25. Schematic diagram of a complete portable radio using transistors in all stages.

Both transistors used as i-f amplifiers are connected in common-emitter circuits, and the collector is connected to a tap on the i-f transformer to avoid loading the transformer. The second i-f amplifier is neutralized by the signal fed through the .1-mfd capacitor C7 back to the secondary of the i-f transformer T3.

The forward bias needed across the emitter-base junction is developed by the resistors in the emitter circuit. R2 develops the voltage needed for the converter, and R6 and R9 develop the voltages for the two i-f stages.

The last i-f transformer T4 is a step-down transformer, and the secondary of this transformer is connected to a diode second detector. This diode rectifies the rf signal so that the audio signal appears in the diode output across the 1500-ohm volume control R12A.

THE AUDIO SECTION

The audio section of this receiver is made up of three stages. A PNP transistor is used in a common-emitter circuit as the first audio frequency amplifier. The circuit used in this stage is similar to the audio amplifier you already studied.

The second audio stage is called the driver stage on the schematic diagram. It is called a driver because it is a power amplifier. It is used to supply power to the push-pull output stage. The push-pull output stage in this receiver uses two transistors in a common-emitter circuit. This stage is a class B push-pull amplifier. It is somewhat different from the class A driver used in the preceding stage. In a class A stage, the average collector current does not change when a signal is applied to the input, but in a class B stage the collector current increases appreciably when the signal is applied. A class B transistor stage, like a class B tube stage, is biased to cut-off. In a transistor circuit this means that there is little or no forward bias across the emitter-base junction.

Notice the resistor R21. The symbol used to identify this resistor is that of a conventional resistor with a circle drawn around it and the letter T inside the circle. This is a thermistor. You will remember that a thermistor is a resistor with a negative temperature coefficient. In other words, if the temperature increases, its resistance decreases. Thermistors are used in this circuit to avoid instability in the output stages which temperature changes would otherwise produce.

SUMMARY

We do not expect you to remember all the circuits used in this receiver. The purpose of presenting this schematic diagram is to give you a general idea of how the various circuits you studied in this lesson may be used together in a complete portable receiver. These circuits are typical of those that you will find in portable radio receivers. In the output stages of some receivers there is only a single transistor instead of the two transistors in the push-pull circuit found in this receiver. You will also find some portables with only one i-f stage. Of course a receiver with a single i-f stage will not have the sensitivity of this receiver.

Spend some time studying this schematic diagram; it will help you become familiar with transistor circuits. This will be a big help to you in later lessons when you study these circuits in more detail.

LOOKING AHEAD

In the next group of lessons you are going to study the various types of circuits found in modern electronic equipment. Your next lesson discusses power supplies.

Power supplies are important because they are found in every piece of electronic equipment. In portable equipment the power supply is usually made up of one or more batteries. In equipment designed to operate from the power line, the power supply will often contain a transformer designed either to step-up or to step-down the power-line voltage. In addition it will contain a rectifier designed to change alternating current to direct current and then some means of smoothing out the pulsating direct current at the output of the rectifier into smooth dc. The power supply will often consist of one or more voltage-divider networks designed to provide more than one operating voltage from a single power supply.

In the lesson on power supplies and in the following lessons we will discuss the circuit in general terms first, pointing out what the circuit must do, what is needed to accomplish the desired results, what the basic circuit looks like, and some of the more important variations of the basic circuit. Then we will go into typical circuits; where possible we will give part values and other pertinent information that might help you become more familiar with these circuits. In the following lessons you will find that you will use many of the basic fundamentals studied in the first thirteen lessons. Do not fail to review these early lessons. It is a good idea to review one or two of your earlier lessons with each new lesson you study. By doing this you will pick up many of the fine points you missed the first time you went over the lesson and, in addition, you will be sure that you do not forget what you have already learned.

ANSWERS TO SELF-TEST QUESTIONS

(a) No. In the common-base circuit the collector current is always less than the emitter current because part of the emitter current leaves the transistor through the base. Although this current is quite small, it does subtract from the emitter current. Therefore the collector current will be less than the emitter current and the current gain of the stage will be less than one.

(b) They are in phase.

(c) Holes.

(d) We can obtain a voltage gain in a common-base amplifier because the output load resistor can be made quite large. Thus, even though the signal current flowing through the load resistor is smaller than the input signal current, the fact that the output load resistor can be made many times the input impedance of the transistor results in the output voltage being greater than the input signal voltage. The output voltage will be the product of the output signal current times the load resistor. The input voltage is equal to the product of the input signal current times the input resistance. As long as the output product is greater, we will have a voltage gain.

(e) A very low input impedance; high output impedance.

(f) To the grounded grid vacuum tube circuit.

(g) Yes. In the common-emitter circuit the signal voltage is applied across the input resistor. None of the emitter current that is flowing to the collector flows through this resistor. The actual signal current flowing in the input is comparatively small. At the same time, the collector current is equal to the emitter current minus any current lost in the base and therefore is much larger than the signal input current. As a result, a current gain is possible with the common-emitter circuit.

(h) The amplified signal voltage is many times the input signal voltage and is 180° out-of-phase with it in the common-emitter circuit.

(i) The input impedance of the common-emitter circuit is much higher than the input impedance of the common-base circuit.

(j) See Fig. 7A for the common-emitter circuit using an NPN circuit and Fig. 7B for a circuit using a PNP transistor.

(k) The common-emitter circuit can be compared with the tube circuit in which a tube is used in a typical grounded-cathode circuit.

(l) The output signal voltage will be in phase with the input signal voltage.

(m) The voltage gain of a common-collector circuit is less than one.

(n) The common-collector circuit has a comparatively high input impedance and a low output impedance. Its input impedance is higher than that of the other two circuits, and at the same time, it is the only circuit in which the output impedance is lower than the input impedance.

(o) The grounded plate or cathode follower circuits.

(p) The common-base circuit and the common-collector circuit.

(q) The common-emitter circuit.

(r) The common-base circuit and the common-emitter circuit.

(s) The common-emitter circuit and the common-collector circuit.

(t) The common-emitter circuit and the common-collector circuit.

(u) The common-collector circuit.

(v) Emitter current I_E; collector current I_C; and base current I_B.

(w) I_{CBO} is the symbol used to represent the collector-base reverse current with the collector-base junction reverse biased and the emitter open.

(x) The symbol V is widely used to represent voltage. The capital C and the capital E have been used to represent the collector and emitter under dc conditions. Therefore, the symbol V_{CE} means the dc collector-to-emitter voltage.

(y) The Greek letter α represents the current gain of a transistor in a common-base circuit.

(z) The maximum value that α can have will be some value slightly less than one. The current gain of the transistor in a common-base circuit will always be less than one. This is due to the fact that the emitter current will be greater than the collector current because some of the carriers leaving the emitter are lost in the base and do not reach the collector.

(aa) The forward current-transfer ratio is the gain of the transistor in a common-base circuit and hence it is equal to α.

(ab) The Greek letter β.

(ac) The alpha cut-off frequency is the frequency at which the current gain of a transistor in a common-base circuit drops to .707 of its gain at lower frequencies.

(ad) It is the frequency at which beta drops to one.

(ae) .96.

We find the value of α by using the formula

$$\alpha = \frac{\beta}{1 + \beta}$$

substituting 24 for β, we get:

$$\alpha = \frac{24}{1 + 24} = \frac{24}{25}$$

$$\begin{array}{r} .96 \\ 25\overline{)24.0} \\ \underline{225} \\ 150 \\ \underline{150} \end{array}$$

$$\alpha = .96$$

(af) Beta equals 99.

We find the value of β by using the formula

$$\beta = \frac{\alpha}{1 - \alpha}$$

substituting .99 for α, we get:

$$\beta = \frac{.99}{1 - .99} = \frac{.99}{.01}$$

$$\beta = \frac{99}{1} = 99$$

(ag) The forward bias across the emitter-base junction is equal to the battery voltage less the voltage drop across R1. There will be a substantial voltage drop across R1 because electrons leaving the base of the transistor flowing through this resistor will develop a voltage drop which subtracts from the battery voltage. Even though the current flowing through the resistor is small, the resistance of the resistor is large so that the net forward bias across the emitter-base junction is only a few tenths of a volt.

(ah) Bias stabilization is necessary because reverse collector-base current due to minority carriers crossing from the collector into the base will effectively reduce the base current. This will reduce the voltage drop across the base resistor and increase the forward bias on the transistor. The increased forward bias may cause sufficient current to flow through the transistor to destroy the transistor.

(ai) The common-base amplifier circuit provides voltage amplification without a phase shift. In some applications, particularly in television receivers, this may be an advantage.

(aj) The low input resistance of a transistor would load a resonant circuit and thus reduce the selectivity of the circuit. By using an untuned secondary and a step-down transformer, the effects of loading on the resonant circuit can be eliminated.

(ak) The collector is connected to a tap on the transformer to eliminate the effect of loading on the tuned circuit, which will cause poor selectivity.

(al) By neutralization - that is, feeding a signal equal to but 180° out-of-phase with the collector-base feedback signal back into the base circuit so that the signal deliberately fed back cancels the signal fed back through the transistor itself.

(am) The oscillator circuit uses a common-base circuit; the signal rf circuit uses a common-emitter circuit. The two circuits are used with the single transistor used in the converter stage.

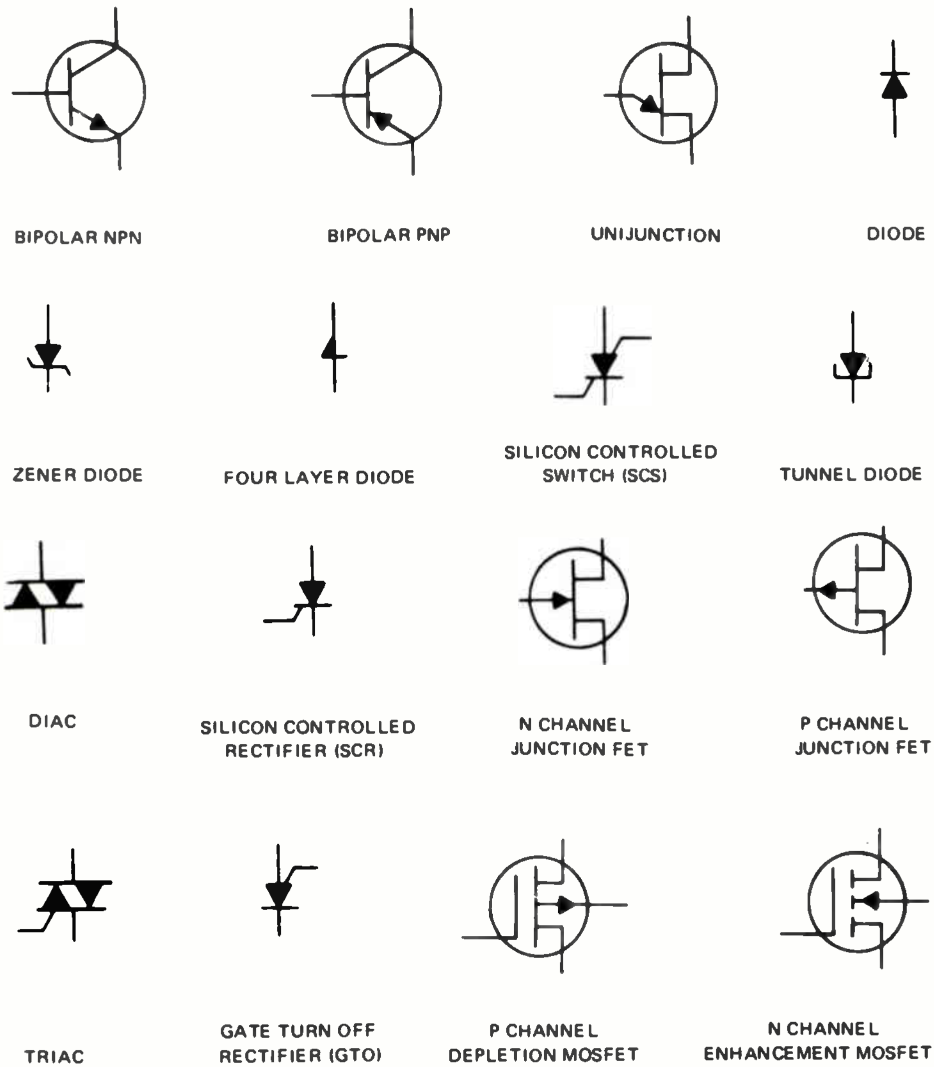


Table I. Semiconductor device symbols.

Lesson Questions

Be sure to number your Answer Sheet B113.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. In a PNP transistor circuit, is the collector made (a) positive or (b) negative with respect to the base?

2. Name the three basic transistor circuits and tell what vacuum-tube circuit each one resembles.

3. Which basic transistor circuit produces a 180° phase shift?

4. Which of the following transistor amplifiers has a current gain? (a) common-base or (b) common-emitter?

5. Which basic transistor circuit has the highest input resistance?

6. (a) What does the term "alpha" applied to a transistor mean? (b) What does the term "beta" mean?

7. If the alpha of a certain transistor is .96, find the beta of the same transistor.

8. If the beta of a certain transistor is 49, find the alpha of the same transistor.

9. In the circuit shown in Fig. 16, the transistor will be very sensitive to temperature changes. How can this be overcome?

10. If a transistor has an alpha cut-off frequency of 100 kc, is it suitable for use as a 455 kc i-f amplifier?

FIRST IMPRESSIONS

First impressions mean a lot in this busy world. An applicant for a job has a pretty tough time making the grade if his appearance and first few words do not make a favorable impression on the employment manager. A salesman likewise gets the "cold shoulder" if there is anything about him which annoys the prospect.

With technical material of any kind, however, first impressions can be very treacherous. Oftentimes a simple technical book will contain a number of apparently complicated diagrams, charts, graphs, sketches or tables. Since we glance mostly at illustrations when inspecting a book, we are apt to get a misleading impression. Also, paragraphs, pages or entire lessons may seem difficult during the first reading, but become almost magically clear during the second or third reading.

If the first impressions of a required task are favorable, fine and dandy; if unfavorable, don't be discouraged, but wade right into the work and give it a chance to prove that first impressions don't always count.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-213

ACHIEVEMENT THROUGH ELECTRONICS

nri

POWER SUPPLIES FOR ELECTRONIC EQUIPMENT

B201

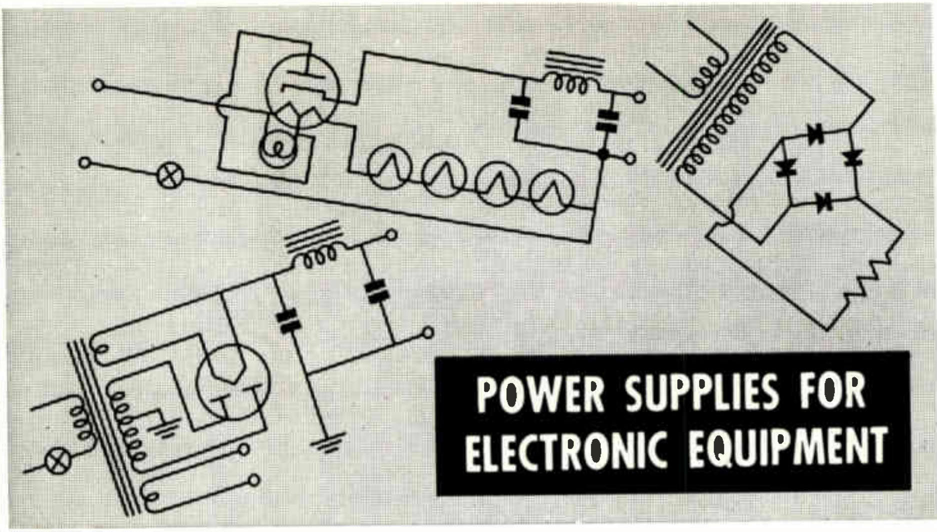
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

POWER SUPPLIES
FOR
ELECTRONIC EQUIPMENT

B201

STUDY SCHEDULE

JDL5M373 1973 EDITION Litho in U.S.A.



POWER SUPPLIES FOR ELECTRONIC EQUIPMENT

One of the most important sections in electronic equipment is the power supply. It is the section that furnishes the operating voltages and currents required by the various stages. If the power supply is not operating properly, the equipment can't do the job it's supposed to do.

In your career as an electronics technician, you will encounter many different types of power supplies. In equipment using tubes you will find power supplies that must supply the heater voltage for the various tubes and, in addition, a dc supply voltage that is often considerably higher than the power line voltage.

On the other hand, the power supply in transistorized equipment will be quite different from that used in tube-operated equipment. Most power supplies in transistorized equipment will have to reduce the voltage to some value less than the line voltage.

However, the current requirements of the power supply in a transistorized piece of equipment may be considerably higher than those of a power supply in a similar tube-operated device. All power supplies have basically the same function, regardless of the parts and circuitry used to make them; the power supply must supply the operating voltages and currents required by the various stages in the equipment.

You have already studied the basic components used in power supplies. In this lesson you will learn more about these components and how they are used together in this particular application. You will be introduced to some new circuits and will learn enough about power supplies to enable you to understand the purpose for which each part in a power supply is used. Once you know why the various parts are used and understand what each one is supposed to do, you should be able to service any power supply defect you encounter.

We will first take up the different rectifier circuits used in modern power supplies. The power supplied by power companies for home and industrial use is ac power, whereas the tubes and transistors used in electronic equipment require dc operating voltages. Therefore, in a power supply designed to operate from a power line, we must have some means of changing the ac to dc. The device used to do this is called a rectifier.

Once the ac is changed to dc by a rectifier, we have what is called a pulsating dc at the output of the rectifier. This is actually dc with ac superimposed on it. A power supply must therefore have some means of filtering or smoothing the pulsating dc to get pure dc. This is done by means of a filter network, which separates the ac and dc components of the pulsating dc at the rectifier output so that only the dc appears at the output of the filter network.

Many power supplies also have some type of voltage-divider network. Such a network is designed to provide several different operating voltages from one power supply. All the tubes or transistors in a piece of electronic equipment may not require the same operating voltage. It is more economical to use a single power supply and a voltage divider than to use a separate power supply for each voltage needed.

The power supplies in modern electronic equipment use solid state rectifiers in most low-voltage applications. Vacuum tubes are seldom used today as rectifiers in such devices as radio or television receivers or in other modern equipment. However, there are still millions of radios and television receivers in use today that do use vacuum-tube rectifiers. Therefore, you will probably have to work on this type of power supply as a service technician even though it is obsolete as far as its use in new equipment is concerned and you still should know how this type of rectifier works. For this reason, we will cover not only the new rectifier circuits using solid state rectifiers, but also a number of the older frequently-used rectifier circuits using vacuum tubes.

Rectifier Circuits

Any device that will pass current in one direction but not in the other direction can be used as a rectifier. You have already seen one example of this type of device: the vacuum tube. In a vacuum tube, as long as the voltage applied to the plate is positive with respect to the voltage applied to the cathode, the current will flow from the cathode to the plate of the tube. However, if the voltage applied to the plate is negative with respect to the voltage applied to the cathode, there will be no current flow through the tube because current cannot normally flow through the tube from the plate to the cathode. Thus, a two-element or diode tube was used for many years as the rectifier in the power supply of radio and television receivers.

The diode tube is entirely satisfactory as a rectifier, but it does have one big disadvantage. In order to handle the currents required in large radio receivers or in television receivers a rectifier tube with a rather heavy cathode or filament is required. Considerable power must be applied to the heater to heat the large cathode or filament, thereby bringing it to the temperature required for it to emit an abundant supply of electrons. Not only does this increase the power consumed by the equipment; also, a substantial amount of heat is given off by the diode and this, in turn, heats up other parts in the equipment. This often contributes to a shortened life of the other parts.

As we mentioned, diode vacuum tubes were used for many years as the rectifiers in radio and television receivers. However, a number of years ago the selenium rectifier began to replace the vacuum tube as the rectifier in entertainment-type equipment.

A typical selenium rectifier is shown in Fig. 1. A selenium rectifier is made up of a series of selenium discs with a coating of selenium oxide on the surface of one side of each disc. Electrons can flow from the selenium to the selenium oxide quite readily, but they cannot readily flow in the other direction, from the selenium oxide to the selenium. Thus, a selenium rectifier permits current to flow through it in one direction, but offers a high resistance to current flow through it in the opposite direction.

This type of rectifier is often called a dry-disc rectifier: "dry" to distinguish it from earlier rectifiers that used a wet chemical solution, "disc" because it is made up of discs. The square plates that are visible in Fig. 1 are cooling fins. The discs used are usually round and are

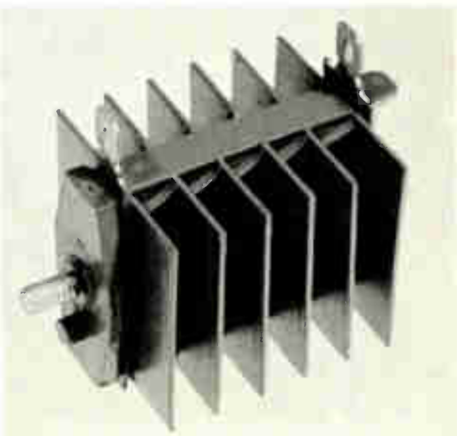


Fig. 1. A typical selenium rectifier designed for use in electronic equipment.

placed between the cooling fins, which are necessary because the rectifier does have some resistance and the current flowing through this resistance produces heat which must be dissipated.

The advantage of the selenium rectifier over the diode tube is that the selenium rectifier does not have a cathode that must be heated, and hence the power required to heat the cathode is saved. In addition, the total heat dissipated into the equipment from the selenium rectifier is somewhat lower than from a tube capable of handling the same current.

Both the vacuum tube rectifier and the selenium rectifier have been replaced in modern radio and television receivers by the silicon rectifier. The silicon rectifier, like the selenium rectifier, does not require any heater power; in addition, a silicon rectifier is much smaller than a selenium rectifier. It has a much lower forward resistance; that is, it offers far less opposition to current flow through it in the forward direction than does a selenium rectifier, and at the same time it has a higher reverse resistance (in other words it will permit a smaller current to flow through it in the reverse direction than a selenium rectifier).

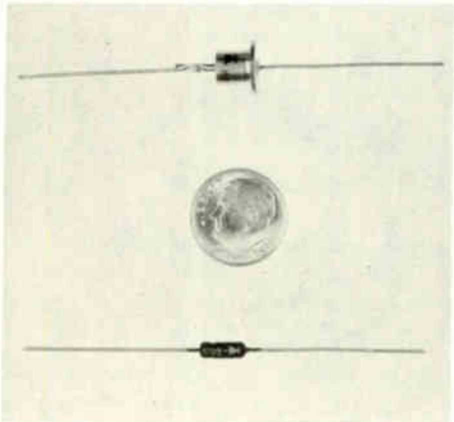


Fig. 2. Two typical silicon rectifiers with a dime between them to show their relative size.

Two typical silicon rectifiers are shown in Fig. 2. The rectifier at the top is called a top-hat rectifier because of its shape. A rectifier of this type and size is capable of handling currents several times those required in a color TV receiver. We have shown a photograph of the two rectifiers with a dime in between them so you can get an idea of the relative size of the two units. The lower rectifier is capable of handling currents of two or three amperes.

In addition to their small size and high current-handling capabilities, silicon rectifiers have another big advantage over selenium and vacuum tube rectifiers due to modern manufacturing techniques: they are relatively inexpensive to manufacture. Furthermore, unless they are overloaded, their life is almost indefinite.

Now, let's see how the various types of rectifiers are used in power supply circuits.

HALF-WAVE RECTIFIERS

You already know that the power supplied by most power companies is ac power and that the voltage supplied has a waveform that is called a sine wave. A typical single cycle is shown in Fig. 3A. A rectifier circuit using a diode tube is shown in Fig. 3B. Let's assume that the waveform shown at A represents the voltage at terminal A of Fig. 3 with respect to the voltage at terminal B. This means that for the first half-cycle (that is, the voltage waveform

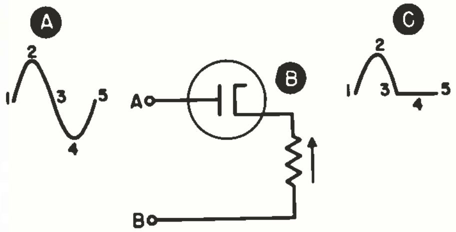


Fig. 3. How current flows in a half-wave rectifier circuit for one ac cycle. We have omitted the rectifier heater to simplify the diagram.

from point 1 to point 3), the plate of the tube will be positive. When the plate of the tube is positive, it will attract electrons from the cathode; therefore, current can flow through the tube.

Thus during the first half-cycle, as the plate voltage starts at point 1 and builds up to point 2, the current through the tube will increase from point 1 to point 2 as shown in Fig. 3C. As the voltage decreases during the first half-cycle from point 2 to point 3, the current through the tube will decrease as shown from point 2 to point 3 in Fig. 3C. When the voltage in Fig. 3A reaches point 3 there will be zero potential between points A and B in the rectifier circuit and current will stop flowing.

During the next half-cycle terminal A will be negative with respect to terminal B. This means that the plate of the tube will be negative; hence no current can flow through the tube. Therefore, the current will be zero (as shown in the waveform in Fig. 3C) as the voltage swings from point 3 to point 5.

The rectifier circuit shown in Fig. 3 conducts when the plate of the tube is positive. Since this occurs during one half the time of each cycle then the rectifier conducts during one half of the cycle but not during the other half. As a result, the rectifier is called a half-wave rectifier. If we operate this type of rectifier from a 60-cycle power line, we will get 60 current pulses through the rectifier during one half-cycle and 60 intervals during which there is no current flow through the rectifier.

Another half-wave rectifier is shown in Fig. 4. Here we have shown a solid-state rectifier in place of the tube. This could be either a selenium rectifier or a silicon rectifier -- the same symbol is used for both types.

Notice the schematic symbol used for the rectifier. Also notice that the arrows indicate that the direction of current flow through the circuit is opposite to the direction in which the arrow points in the schematic symbol. The reason for this is that in the early days of electricity, scientists thought that current flowed from positive to negative. Therefore, this symbol was designed to show the direction in which current flowed. But it was discovered later that current flow was actually electron flow and that it flowed from negative to positive, a direction opposite from that in which the early scientists thought it flowed. Although we know current flow is from negative to positive, we still use the same symbol; it has never been changed so that the arrow is actually pointing in the direction opposite to the direction of electron flow.

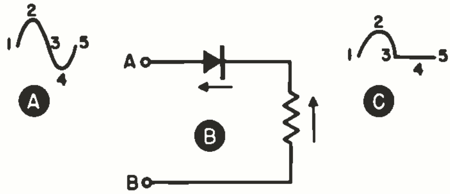


Fig. 4. A half-wave selenium rectifier circuit.

In a power supply of this type using a selenium rectifier, with terminal A positive with respect to terminal B, the selenium rectifier will offer only a low resistance to the flow of current through it; therefore, current flows in the circuit from B through the load and through the rectifier and back to terminal A. During the next half-cycle, when terminal A is negative with respect to terminal B, the selenium rectifier will offer a very high opposition to the flow of current through it so that there will be little or no current flow through the load (as shown in Fig. 4C).

We mentioned that the schematic symbol for the rectifier in Fig. 4 also represents a silicon rectifier. If a silicon rectifier is used, the rectifier will simply consist of a PN junction. The P-type material will be on the side represented by the arrow and the N-type material by the flat line. With a PN junction rectifier in the circuit when terminal A is positive and terminal B is negative, we will have a positive voltage applied to the P side of the junction and a negative voltage applied to the N side of the junction. The negative voltage will repel electrons from the N side of the junction across the junction into the P-type material. Electrons will be attracted through the P-type material by the positive potential applied to it. In other words, there will be a forward bias placed across the junction and current can readily flow through the rectifier because the carriers can cross the junction.

During the next half-cycle when the polarity reverses, terminal A will be negative and terminal B positive. Thus there will be a negative voltage, applied to the P side of the junction, which will repel electrons and prevent them from crossing the junction. At the same time there will be a positive voltage, applied on the N side of the junction, which will prevent any holes from the P side crossing the junction. In other words, there will be a reverse bias placed across the junction; hence the carriers cannot cross the junction and there will be no current flow through the circuit.

Of the three types of half-wave rectifiers, the silicon-type rectifier is the most widely used in modern equipment because of its small size, low cost and very low forward voltage drop.

Sometimes, in order to operate a piece of electronic equipment, a higher voltage is required than can be obtained directly from the power line. Under these circumstances a step-up transformer may be used to step up the voltage as shown in Fig. 5. The secondary-to-primary turns-ratio is simply adjusted to provide the required voltage step-up. A half-wave rectifier is then used as shown to rectify the ac and change it to pulsating dc. The operation of the half-wave rectifier is exactly the same in the circuit as

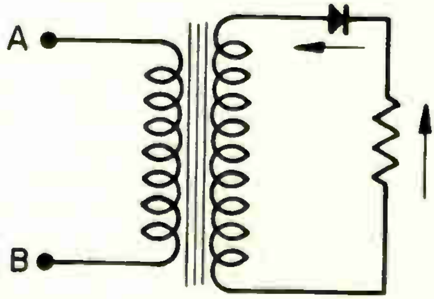


Fig. 5. A half-wave rectifier circuit using a power transformer to produce an output voltage greater than the power line voltage.

it was in the preceding circuits; however, the rectifier will have to have a higher voltage rating to make up for the fact that it is being used in a higher voltage circuit.

The disadvantage of the circuit shown in Fig. 5 is that power transformers are comparatively expensive. By means of the circuit shown in Fig. 6 a voltage approximately twice the voltage obtainable from the half-wave rectifier circuits shown in Figs. 3 and 4 can be obtained. This circuit can be operated directly from the power line and is known as a voltage-doubler circuit.

The operation of this circuit is quite simple. During one half-cycle terminal A will be negative with respect to terminal B. During this half-cycle electrons flow from A into the side of the capacitor C1 marked with the minus sign. The electrons flowing into this side of the capacitor force electrons out of the other side leaving a positive charge on this side of C1. The electrons leaving the positive side of C1 flow through the rectifier D1, back to side B of the power line which is positive and which will attract electrons. Thus, during this half-cycle, when terminal A is negative with respect to terminal B, capacitor C1 is charged with the polarity shown. The peak charge on C1 will be equal to the peak value of the ac input voltage.

During the next half-cycle, when terminal A is positive with respect to terminal B, we have a situation where the voltage between terminals A and B is effectively placed in series with the voltage charging capacitor C1. These series-connected voltages will cause a current to flow through the load and through D2. During this half-cycle terminal A is positive with respect to terminal B. This means that terminal B is negative. Electrons flow from terminal B through the load and through D2. They are attracted by a positive voltage which is equal to the voltage across C1 plus the line voltage. Thus the peak voltage that can be developed across the load will be equal to twice the peak line voltage.

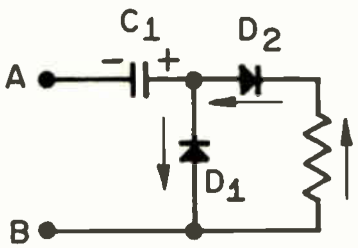


Fig. 6. A half-wave voltage-doubler circuit.

You might wonder why the current flows through only one rectifier during each half-cycle. During the first half-cycle, when terminal A is negative with respect to terminal B, the electrons flowing through D1 and charging C1 cannot flow through D2 because the diode is connected in such a way as to prevent current flow through it in that direction. Similarly, during the next half-cycle, when terminal B is negative and terminal A is positive, current cannot flow through D1 because it would have to flow through it in the reverse direction. Current can flow through the diodes only in the direction shown and it will flow through D1 during one half-cycle and through D2 during the other half-cycle.

This type of power supply is known as a half-wave doubler circuit. It is called a voltage-doubler circuit because the voltage across the load is effectively double the line voltage.

It is called a half-wave circuit because there is a current pulse to the load during only one half of each cycle. The half-wave voltage-doubler circuit is widely used in modern radio and television receivers. It's a very important circuit and you should be sure you understand how it works before leaving it.

When a half-wave rectifier is used in the power supply, the current will flow through the rectifier in a series of pulses. With a 60-cycle power supply line, there will be 60 pulses per second: one pulse during each positive-half cycle and nothing during each negative-half cycle. The net result is that you will have current flowing through the rectifier for no more than half the time. This results in a pulsating dc output from the rectifier that is rather difficult to smooth out to the pure dc required in most equipment to operate the tubes and/or transistors. A somewhat better arrangement is the full-wave rectifier that passes current during both halves of the ac voltage cycles.

FULL-WAVE RECTIFIERS

A typical full-wave rectifier circuit is shown in Fig. 7. Notice that the tube used is a twin diode tube.

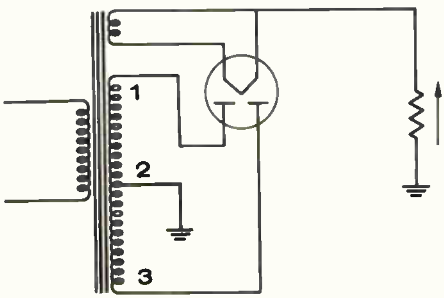


Fig. 7. A full-wave rectifier circuit using a single rectifier tube with two plates.

The tube has two plates and a single filament which is used with both plates. In operation, this tube acts as two separate diode tubes.

The power transformer used in the rectifier circuit has three windings. The primary winding is the winding that connects to the power line. A low-voltage winding is used to provide the current required to heat the filament of the rectifier tube. It serves no other purpose as far as the operation of the rectifier circuit is concerned. This winding is often referred to as the filament winding.

The high-voltage winding on the transformer is the winding that will supply the pulsating current to the load resistor. Notice that this winding has a center tap. In operation, one half of the winding first supplies the current and then, during the next half-cycle, the other half of the winding supplies current to the load.

We can see how this rectifier circuit works if we consider one half-cycle during which terminal 1 of the high-voltage secondary is positive. This means that terminal 2 will be negative with respect to terminal 1 and terminal 3 will be even more negative. Electrons will leave the center tap, terminal 2, and flow through ground to the load resistor. They will flow through the load resistor to the filament of the rectifier tube and then be attracted to the plate connected to terminal 1 because this plate has a positive voltage applied to it. No electrons will flow to the other plate because this plate is negative with respect to both terminals 1 and 2.

During the next half-cycle, the polarity of the secondary voltage will reverse. At this time terminal 3 will be positive, terminal 2 negative with respect to it, and terminal 1 even more negative. During this half-cycle electrons will leave terminal 2 and flow through ground to the load, through the load and to the filament of the rectifier tube, and then to the plate connected to terminal 3 because this plate now has the positive voltage applied to it. No electrons will flow to terminal 1 because terminal 1 is negative with respect to both terminals 2 and 3.

Notice that with the full-wave rectifier circuit we get a current pulse through the load resistor during each half-cycle. This means that for a 60-cycle power line we will get 120 pulses of current through the load. Since there is current flowing through the load during each half-cycle, this type of rectifier produces an output that is much easier to filter to a smooth dc than the output from a half-wave rectifier.

Either selenium rectifiers or silicon rectifiers can be substituted in this type of circuit in place of the rectifier tube. This type of circuit was widely used in older television receivers along with vacuum tubes. Modern TV receivers use silicon rectifiers and frequently use bridge-rectifier circuits or voltage-doubler circuits in place of this circuit.

BRIDGE-RECTIFIER CIRCUITS

One of the disadvantages of the full-wave rectifier circuit shown in Fig. 7 is that it requires a transformer with a center-tapped secondary. The total voltage across the entire secondary winding is actually twice the voltage between the center tap and either end of the secondary winding. This type of transformer is more expensive to manufacture than a transformer without a center tap

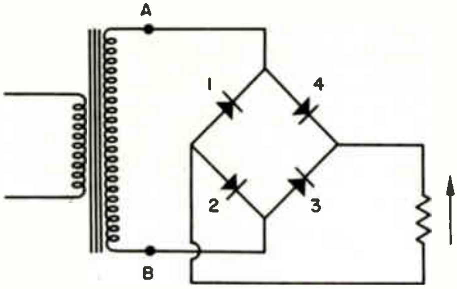


Fig. 8. A bridge-rectifier circuit.

because twice as many turns are required on the secondary to get the required voltage.

A circuit that gets around the requirement of a center-tapped secondary is shown in Fig. 8. This is called a bridge-rectifier circuit; it is also often called a full-wave bridge-rectifier circuit because current flows to the load during each half-cycle.

A quick look at the circuit immediately shows us that four rectifiers are required in a circuit of this type. At one time this was a disadvantage because of the cost of rectifiers, but silicon rectifiers are comparatively inexpensive today and it is usually more economical to use the extra two silicon rectifiers and avoid the center tap on the secondary winding of the power transformer. The power transformer shown in Fig. 8 would be far more economical to manufacture than the one shown in Fig. 7.

The operation of the bridge rectifier is comparatively simple. When terminal A is positive and terminal B is negative, current will flow from terminal B through the rectifier marked 2 on the diagram and then through the load to the junction of rectifiers 3 and 4. It will then flow through rectifier 4 back to terminal A on the transformer. During the next half-cycle, when terminal A is negative and terminal B is positive, current will flow from terminal A on the transformer through the rectifier marked 1 on the diagram and then through the load back to the junction of rectifiers 3 and 4. This time the current will flow through rectifier 3 back to terminal B of the power transformer.

Notice that during each half-cycle current flows through two of the rectifiers. During one half-cycle it will flow through rectifiers 2 and 4 and during the other half-cycle it will flow through rectifiers 1 and 3. Also notice that current flows during both half-cycles; therefore, this bridge rectifier is a full-wave rectifier.

Bridge rectifiers have been used in television receivers where comparatively high operating voltages and high currents are required. The bridge circuit eliminates the need for the center tap on the power transformer secondary winding and thus reduces the cost of the transformer. The voltage regulation (the ratio of the full-load voltage to the no-load voltage) obtainable with this type of power supply is as good as the regulation that can be obtained from the full-wave rectifier circuit shown in Fig. 7.

FULL-WAVE VOLTAGE DOUBLERS

A full-wave voltage-doubler circuit is shown in Fig. 9. One advantage of this type of circuit is that we get the same load voltage as you have in a circuit like the bridge rectifier shown in Fig. 8, although only half as many turns are required on the secondary of the power transformer. This will result in a savings in the cost of the power transformer. Another advantage of this type of circuit is that only two rectifiers are required instead of the four required in the bridge-rectifier circuit.

The operation of the full-wave voltage-doubler circuit is quite simple. During one half-cycle terminal A of the power transformer secondary will be negative and terminal B will be positive. During this half-cycle current flows from terminal A through diode D2 to the capacitor C2 charging the capacitor as shown. Electrons flow into the negative side of the capacitor and out the positive side back to terminal B of the power transformer. During the next half-cycle, terminal B of the power transformer secondary will be negative and terminal A will be positive. During this half-cycle electrons leave terminal B of the power transformer and flow into capacitor C1. They flow into the side of the capacitor marked with the minus sign and force electrons out of the plus side. Electrons leaving the plus side flow through diode D1 back to terminal A of the power transformer, which is positive.

The capacitors C1 and C2 are connected in series and they supply the voltage to the load. The capacitors are charged by the current flowing

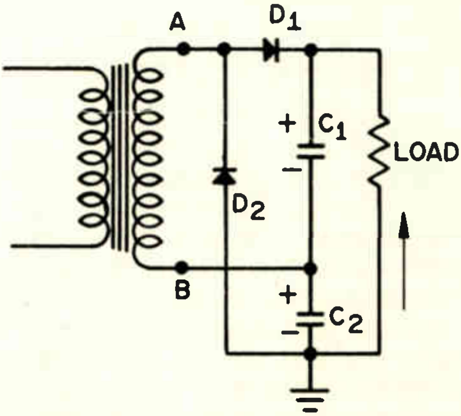


Fig. 9. A full-wave voltage doubler.

through the diodes, and since there is a charging pulse each half-cycle we will get 120 charging pulses from a 60-cycle power line. Since there is a charging pulse during each half-cycle, the circuit is a full-wave rectifier. The actual voltage that will be available across C1 and C2 in series will depend upon the power transformer secondary resistance, the resistance of the diodes when they are conducting, the size of the two capacitors, and the size of the load. As the resistance of the load increases, the current that will flow through the load decreases, and the charge on each capacitor becomes closer to the peak value of the ac voltage between terminals A and B.

Notice that in the diagram we have shown the direction of current flow through the load. The load current is supplied entirely by the charged capacitors C1 and C2. As the capacitors supply current to the load, electrons leave the negative plate of C2 and flow through the load in the direction shown. These electrons flow into the positive side of C1 forcing electrons out of the negative side into the positive side of C2. Thus the current flow through the load tends to reduce the charge across the capacitors. Of course, during each half-cycle one of the diodes conducts to build the charge across one of the capacitors up towards the peak value of the line voltage.

While this type of rectifier circuit offers some advantages over the circuits shown in Figs. 7 and 8, it does not have as good voltage regulation as they have. However, by using capacitors of large capacity for C1 and C2, and with modern silicon rectifiers that have a very low resistance when they are conducting, reasonably good voltage regulation can be obtained from this type of power supply. You will find the power supply widely used both in monochrome and color television receivers. Be sure you understand how it operates because you will run into it frequently.

In comparing this power supply with the half-wave doubler circuit shown in Fig. 6, we immediately see that the full-wave doubler circuit is best suited to equipment where a power transformer is used. The half-wave voltage-doubler circuit is widely used in equipment where no power transformer is used.

SUMMARY

The rectifier circuits that we have discussed in this section of the lesson are extremely important. As you know, the power supplied by the power companies is alternating current and the tubes and transistors in electronic equipment require direct current for their operation. Therefore, equipment designed to operate from the power line must use some type of rectifier to convert the alternating current to direct current. One of the circuits shown in this section of the lesson is likely to be found in any type of electronic equipment you will service.

The half-wave rectifier circuit shown in Fig. 5 is perhaps the most widely used. All table model radio receivers use this type of rectifier circuit without a power transformer so that the receiver can operate directly from the power line. The half-wave voltage-doubler circuit shown in Fig. 6 is widely used in television receivers where operating voltages higher than those that can be obtained directly from the power line are needed. In some of the older radio receivers and many older television receivers a full-wave rectifier circuit such as shown in Fig. 7 will be found. This type of circuit was used almost exclusively in television receivers before the development of low-cost selenium and silicon rectifiers.

The bridge-rectifier circuit shown in Fig. 8 is used in many television receivers, particularly large television receivers where fairly high voltages and high currents are required. This type of circuit is also used in some transistorized equipment where lower than line voltages are required. In this instance, instead of being a step-up transformer, the transformer will be a step-down transformer that steps the line voltage down to a low value. The bridge-rectifier circuit is then used so that good voltage regulation can be obtained and so that at the same time a comparatively large current can be taken from the power supply.

The full-wave voltage-doubler circuit has been used in many television receivers -- in those which use power transformers where voltages higher than the line voltage are required -- yet at the same time, the transformer serves primarily as an isolation transformer. In other words, the secondary voltage is approximately equal to the primary voltage. The higher voltage required is obtained by the voltage-doubling action. The output from the full-wave doubler circuit is somewhat easier to filter or smooth out than the output from the half-wave doubler circuit, because with the former we have 120 pulses per second through the load and with the latter we have only 60 pulses per second.

Before leaving this section regarding rectifier circuits, why not try to draw the circuits yourself? It would be worthwhile, and you don't have to draw them from memory -- copying them first from the book will help you to remember what they look like. Eventually, you'll be able to draw them from memory and recognize various circuits on the schematic diagram of any radio or TV receiver or piece of electronic equipment you may encounter.

After reviewing this section, do the following self-test questions.

SELF-TEST QUESTIONS

(a) In a half-wave rectifier circuit such as the circuit in Fig. 5, how many current pulses per second will there be through the load when the power line frequency is 60 pulses per second?

(b) What is the disadvantage of a half-wave rectifier circuit?

(c) What is the purpose of the diode marked D1 in the half-wave voltage-doubler circuit shown in Fig. 6?

(d) Why is the circuit shown in Fig. 7 called a full-wave rectifier circuit?

(e) What is the disadvantage of the full-wave rectifier circuit shown in Fig. 7?

(f) What are the advantages of the bridge-rectifier circuit shown in Fig. 8?

(g) What advantage does the voltage-doubler circuit shown in Fig. 9 have over the voltage-doubler circuit shown in Fig. 6?

(h) What are the advantages and disadvantages of the full-wave voltage doubler over the bridge-rectifier circuit?

Filter Circuits

The output from the rectifiers we discussed in the preceding section is not pure dc. Instead, it is pulsating dc: "direct" because it flows in only one direction, "pulsating" because it is varying in amplitude rather than flowing steadily. A pulsating dc voltage is a voltage that does not change polarity, but does change in amplitude. The voltage at the output of a half-wave rectifier will be zero during one half of each cycle and swing in a positive direction during the other half of each cycle.

Looking at the half-wave circuits shown in Fig. 3 and 4, you can consider the rectifier more or less as a switch. During one half-cycle the switch is closed so that the load is connected directly across the power line, and during the other half-cycle the switch is open so that no voltage is applied to the load. The voltage across the load in a half-wave rectifier circuit looks like Fig. 10A. The first half-cycle represents the cycle when the switch is closed and the load is connected directly across the power line, and the second half-cycle represents the cycle when the switch is open and there is no voltage applied across the load. We have shown what the voltage across the load will look like for four cycles in Fig. 10A.

In a full-wave rectifier circuit such as shown in Fig. 7, you have two switches. During one half-cycle one switch closes and connects the load across one half of the power transformer secondary; then, during the next half-cycle, the other switch closes and connects the load across the other half of the transformer secondary.

In the bridge rectifier shown in Fig. 8 two rectifiers act as switches and close to connect the load across the transformer secondary during one half-cycle; then during the next half-cycle the other two switches close, giving the effect of turning the load around so that the current flows through it in the same direction and the voltage applied across the load has the same polarity. The output from a full-wave rectifier circuit will produce a voltage across the load that looks like Fig. 10B.

As you can see from the waveforms shown in Fig. 10, the output taken directly from the rectifier is not a pure dc. There is a voltage, but the voltage drops to zero, builds up to the maximum value, and drops to zero again. In the half-wave circuit it remains at zero for a half-cycle and then builds up again in a positive direction. In the full-wave rectifier circuit the voltage builds up across the load during each half-cycle. In either case, this pulsating voltage will cause a pulsating current through the load which is entirely unsuitable for use in electronic equipment. Fortunately, there are convenient methods that can be used to filter or smooth this voltage to a pure dc voltage.

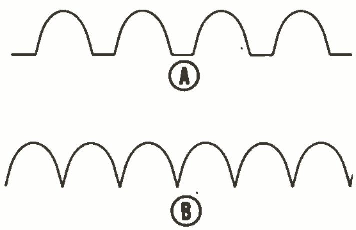


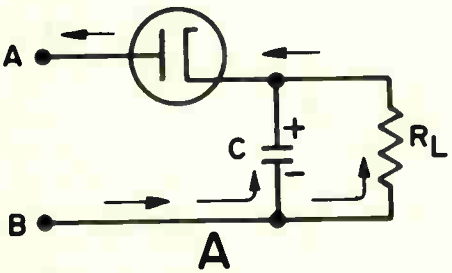
Fig. 10. Output voltage from half-wave and full-wave rectifier.

The pulsating dc voltage at the output of the rectifier is actually a dc voltage with an ac voltage, called a ripple voltage or a hum voltage, superimposed on it. The circuits used to get rid of this ripple or hum voltage are called filter circuits. There are a number of different types of filter circuits found in electronic equipment; in this section we will cover some of the circuits more commonly used.

THE SIMPLE CAPACITOR CIRCUIT

One of the simplest filters is the single capacitor filter shown in Fig. 11. In Fig. 11A we have shown a rectifier circuit using a tube and in Fig. 11B a rectifier circuit using a silicon rectifier. Notice that the circuits are practically identical -- we simply changed the rectifying devices in the two circuits.

The simple capacitor-type filter is sometimes used in circuits where the current drained or taken from the power supply is low. If the rectifier must supply high current to the circuit, this type of filter is generally unsatisfactory because there will be too much ripple or hum present across the load. In other words, the simple filter is simply not capable of eliminating all the ac or ripple voltage present at the output of the rectifier.



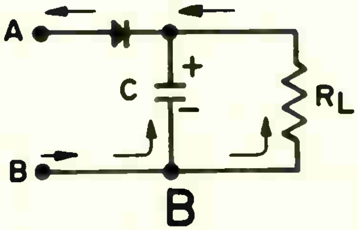


Fig. 11. A simple capacitor-type filter.

Both circuits shown in Fig. 11 work in the same way. Current flows through the rectifier during one half-cycle, as in the half-wave rectifier circuits we studied previously. When terminal A is positive, electrons will flow from terminal B through the load and through the rectifier back to terminal A. At the same time, electrons will flow into the negative side of the capacitor and out the positive side and through the tube or silicon rectifier back to terminal A. The capacitor eventually will be charged to a value almost equal to the peak line voltage. This will happen when the ac line voltage reaches its peak value with terminal A at its peak positive voltage with respect to terminal B.

Now, if the load on the rectifier circuit is light (that is, if the load resistor is a high resistance that draws very little current), as the ac input voltage between terminal A and B drops, capacitor C will begin to supply the current required by the load. Electrons will start to leave the negative side of the capacitor and flow through the load resistor back to the positive side of the capacitor. They will continue doing this as the ac voltage drops to zero and remains at zero during the next half-cycle and starts to build up again in the positive direction. The capacitor will continue to supply current to the load as long as the voltage across the capacitor is greater than the ac input voltage. Eventually, the input voltage will reach a value greater than the capacitor voltage; then we'll get a current flow into the capacitor and through the rectifier to recharge the capacitor.

In Fig. 12 we have shown the ac input voltage. Notice that during the first half-cycle between points 1 and 2 in Fig. 12A the ac voltage is increasing in a positive direction. Let's assume that terminal A in Fig. 11 is becoming positive with respect to terminal B. This explanation applies to both of the two circuits shown. During this first half-cycle the capacitor is charging and follows a curve as shown from point 1 to point 2 in Fig. 12B. Now, as the ac cycle drops from point 2 to point 3 on curve A in Fig. 12, the voltage drops faster than the capacitor discharges. During the interval from point 2 to point 5 and almost to point 6, as shown in Fig. 12A, the capacitor discharges very little. The discharge is shown on the curve from point 2 over to the number 5 on curve B. At this point the ac input voltage exceeds the capacitor voltage, so the capacitor is recharged again.

In circuits where the drain or load current is low, the capacitor will discharge very little between current pulses that recharge it so that the voltage across the capacitor and hence the voltage across the load remain almost constant. Of course, as the requirements of the load increase, the capacitor will discharge more so that there will be more of a voltage drop across the capacitor and across the load than in circuits where current drain is low.

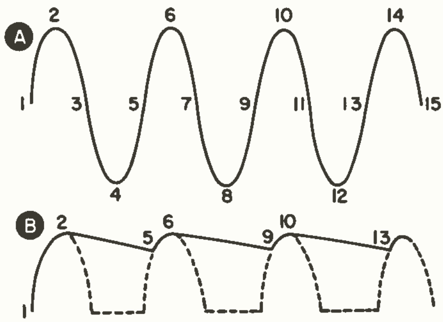


Fig. 12. Voltage waveshapes for a simple capacitor filter. (A) Input voltage; (B) output voltage.

There are several important points you should notice in the circuits shown in Fig. 11. Notice that the current does not flow during an entire half-cycle, but flows only when the line voltage exceeds the voltage across the capacitor. This may be for a very short interval if a load is a high resistance and draws very little current, or it may be for a sizable portion of a half-cycle if the load is a low resistance and draws a high current from the power supply. However, since the current flows through the rectifier in pulses, then the current pulse through the rectifier must be many times the average dc current flowing through the load. This is because the pulse or current that flows through the rectifier during the interval in which the rectifier is conducting must supply enough current to the capacitor to charge the capacitor and make up for the current it is going to supply for the remainder of the cycle.

When the rectifier is not conducting it is because the voltage across it is what we call a reverse voltage. In other words, it has a polarity opposite from that which the rectifier needs to conduct. In the case of the vacuum tube circuit this means that the plate of the tube is negative with respect to the cathode; in the case of the silicon rectifier, that there is a reverse bias across the junction.

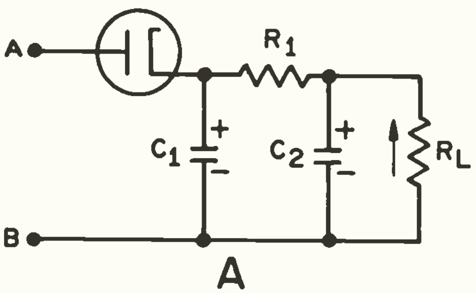
One of the important characteristics of a rectifier is the maximum peak reverse voltage that can be placed across the rectifier before it breaks down. In the circuit shown in Fig. 11 the capacitor will be charged as shown and the charge can equal the peak line voltage. During the next half-cycle, when the polarity of the input voltage reverses, terminal A will be negative and terminal B will be positive. When this voltage reaches its peak, the peak reverse voltage across the rectifier will be equal to twice the peak line voltage. The rectifier must be able to withstand this voltage without breaking down. This important characteristic, by which rectifiers are rated, is usually referred to as "PRV" (peak reverse voltage), although it may also be called "PIV" (peak inverse voltage). The two are simply the maximum reverse or inverse voltages that can be applied across the rectifier without its breaking down. In circuits such as those shown in Fig. 11, the PIV should be considerably higher than twice the peak line voltage in order to allow a reasonable safety factor.

As we mentioned previously, the simple capacitor-type filter shown in Fig. 11 is usable only where a small current is required by the load. If the current required is small, the output capacitor can be made large enough so that it discharges very little between pulses. If the current required by the load is high, on the other hand, then the capacitor will discharge appreciably between charging pulses, resulting in a varying voltage applied to the load. This is essentially the same as applying dc mixed with ac to the load. Additional filtering is required in applications of this type in order to eliminate ac so that we will have pure dc across the load.

AN R-C FILTER

An improved filter, which is often called a pi filter because it looks like the Greek letter pi (π), is shown in Fig. 13. You will notice that this filter consists of two capacitors, C1 and C2, and a filter resistor, R1.

The operation of the half-wave rectifier and capacitor C1, which is called the input filter capacitor, is the same as in the simple capacitor filter shown in Fig. 11. The rectifier tube passes current pulses to charge capacitor C1 with the polarity indicated on the diagram. However, if the



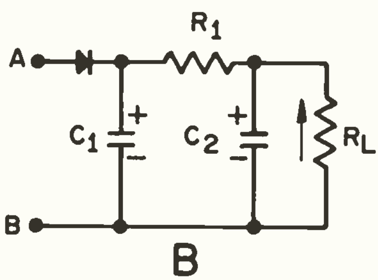


Fig. 13. An R-C pi-type filter.

load resistance R_L is low enough to draw appreciable current from the supply, then the voltage across C1 will discharge appreciably during the portion of the cycle when the rectifier tube is not conducting. Thus, we have dc with an ac superimposed on it across C1.

Now, to see the action of R1 and C2, let us first consider how the capacitor C2 reacts to ac and to dc. Remember that a capacitor is a device that will not pass dc -- it can be charged so that a dc voltage will exist across it, but applying dc to the plates of the capacitor will not cause a current to flow through it. Although electrons cannot cross through the dielectric of a capacitor, however, applying ac to the dielectric of a capacitor yields the effect of a current flowing through it. This is due to the fact that electrons will flow first into one plate and then into the other as the polarity of the ac voltage reverses.

You will remember from your study of capacitors that a capacitor offers what is called capacitive reactance (or opposition) to the flow of ac through it. The exact reactance that any capacitor will offer to the flow of ac through it is given by the formula:

$$Xc = \frac{1}{6.28 \times f \times C}$$

You can see from this formula that the larger the value of the capacitor, the lower the capacitive reactance will be to an ac voltage of a particular frequency. In Fig. 14A we have shown how the filter consisting of R1 and C2 reacts to dc; in Fig. 14B, we have shown how it reacts to ac.

As you know, a "perfect" capacitor will not pass dc -- although no

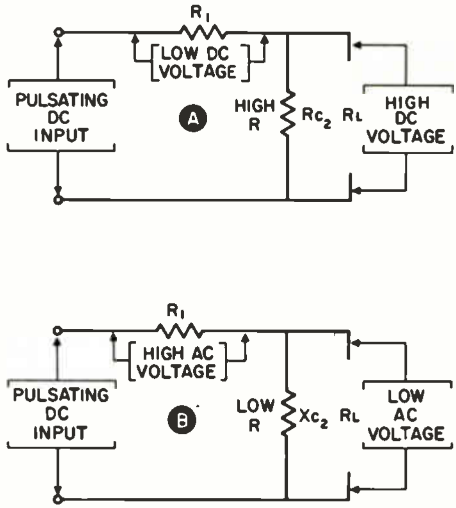


Fig. 14. Equivalent circuits showing the reaction of an R-C filter to dc at A, and to ac at B.

capacitor is really perfect because there will always be a small dc current (called a leakage current). In other words, the capacitor offers a very high resistance to dc, so we have shown it as resistor R_{c2}. Most of the dc voltage applied to the input of this filter network will appear across the capacitor, and there will be very little dc dropped across the resistor R1. The exact drop across this resistor will depend upon the size of the resistor and the current drawn by the load.

Now look at Fig. 14B which shows the reaction of the circuit to an ac voltage. The capacitor has a very low reactance to ac, so most of the ac voltage will be dropped across R1, because the resistance of R1 is much higher than the reactance (X) of C2. R1 and C2 act as a voltage divider network with most of the ac being dropped across R1 because its resistance is much higher than the reactance of C2.

There is another way of looking at this type of power supply which may help you see exactly what is happening in the circuit. Refer back to Fig. 13; the explanation applies to the circuit using the vacuum tube rectifier shown at A and to the one using the silicon rectifier shown at B.

When terminal A is positive with respect to terminal B, the diode will conduct and current can flow through the diode and through the load. During this part of the cycle electrons will flow into the plates of C1 and C2 marked with a minus sign. At the same time electrons will flow out of the other plate of both capacitors. Electrons leaving the positive side of C1 will flow directly through the diode being attracted by the positive voltage at terminal A. Because the resistance of the rectifier is low, C1 can charge up to a value almost equal to the peak ac voltage. However, the electrons leaving the positive side of C2 must flow through the filter resistor R1. Thus, capacitor C2 cannot charge to as high a voltage as capacitor C1.

When the ac input voltage drops so that the diode no longer conducts, capacitors C2 and C1 begin supplying the power required by the load. However, capacitor C1 is charged to a higher voltage than capacitor C2. Hence, capacitor C1 begins supplying power to the load and also tries to charge capacitor C2. Since electrons flowing from the negative side of C1 to the positive side must flow through filter resistor R1, the attempt of these electrons to charge C2 and flow through the load resistor will be somewhat restricted by R1. The net effect is that the resistor R1 prevents C2 from charging to as high a peak voltage as it would if R1 were not on the circuit. Because C1 charges to a higher voltage than C2, and while discharging tends to charge C2, the voltage across C2 is more nearly constant than the voltage across C1.

AN L-C FILTER

The disadvantage of the resistor-capacitor type of filter shown in Fig. 13 soon becomes apparent if you consider the size of the resistor and capacitor needed to obtain effective filtering and also the effect that the high value of resistance in the circuit has on the dc voltage present.

Consider Fig. 13 again for a minute. Suppose that the ac component of the pulsating dc across C1 is 10 volts and that the maximum ac component that can be applied to the load is only 1 volt. This means that the filter network consisting of R1 and C2 must produce a 9-to-1 voltage division. In other words, of the 10 volts ac appearing across C1 we must drop 9 volts across R1 and 1 volt across C2. This means that the resistance of R1 must be about 9 or 10 times the reactance of C2.

A 25-mfd capacitor (which is fairly large, particularly if it must be built to withstand high voltages), has a reactance of about 100 ohms at a frequency of 60 cycles. If we used such a capacitor for C2, then the resistance of R1 would have to be 10 times its reactance, or about 1000 ohms. If the current drawn by the load is 100 milliamperes, then the voltage drop across R1 due to the load current flowing through it will be 1000 ohms times .1 amp (100 ma), or 100 volts. This means that we will be losing 100 volts of our dc voltage across R1. Furthermore, the power being wasted by this resistor will be

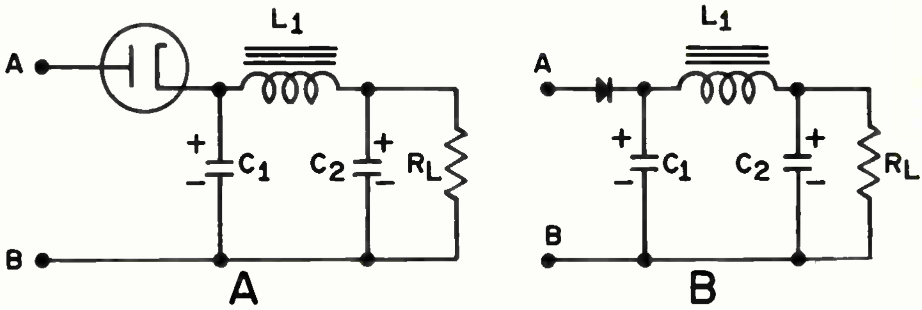


Fig. 15. An L-C filter.

equal to the voltage across it times the current flowing through it, which is 100 × .1, or 10 watts. You can see that we have an appreciable voltage drop and a sizable amount of power wasted by this resistor.

In some cases the current drawn through the filter resistor is not so high that the voltage drop across the resistor cannot be tolerated. Thus you will see this type of filter used in equipment when the current taken from the power supply is moderate. In equipment when the current drawn is high, a different type of filter is used. Such a filter is shown in Fig. 15. Notice that this circuit is identical to Fig. 13, except that we have substituted an iron-core choke for R1. The action of this filter network is quite similar to that of the filter network shown in Fig. 13. Again, capacitor C1 is charged by the rectifier and because there is no resistance in the circuit other than the rectifier resistance, it charges to a fairly high value. Just as before, however, the voltage across it will be pulsating (the equivalent of dc with ac superimposed on it).

Now consider the reaction of the choke to ac and dc. You will remember that a choke has inductance, and an inductance offers reactance to the flow of ac through it. At the same time the only opposition that the choke will offer to the flow of dc through it is due to the resistance of the wire used to wind the coil. By using a large size wire, this resistance can be kept quite low. The choke may have a dc resistance of 100 ohms or less and at the same time have a reactance of several thousand ohms to the 60-cycle ac applied to it.

The reaction of the L-C filter to dc is shown in Fig. 16A, and the reaction to ac is shown in Fig. 16B. In 16A we see that as far as the dc is concerned, the choke acts as a

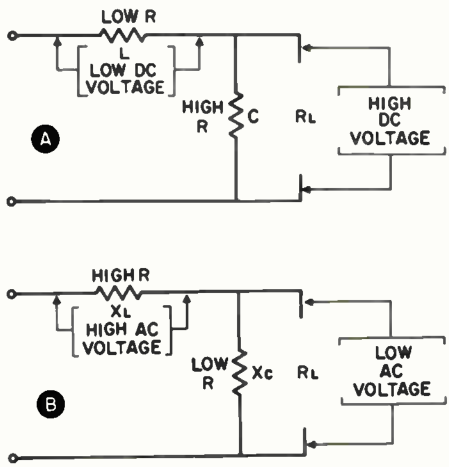


Fig. 16. Equivalent circuit showing the reaction of an L-C filter to dc at A, and to ac at B.

low resistance while the capacitor acts as a very high resistance. Thus, we have practically all of our dc voltage appearing across the capacitor and very little of it being lost across the choke. However, as shown in 16B, the choke acts as a high resistance to the ac while the capacitor acts as a low resistance. Thus, most of the ac voltage will appear across the choke and very little of it will appear across the capacitor.

There is another way of looking at the action of the L-C filter. We can consider the charging of the capacitors and the opposition offered by the choke more or less in the same way as we considered the action of the R-C filter in Fig. 13. During the first half-cycle, when terminal A of either rectifier circuit is positive and terminal B is negative, the rectifier will conduct. Electrons will flow from terminal B into the sides of C1 and C2 marked with a minus sign. These electrons will force electrons out of the positive side of C1 and they will flow through the rectifier with little or no opposition. However, the electrons leaving the positive side of C2 will encounter opposition in the choke. This is because at the instant when the electrons first try to get through the choke there is no magnetic field built up in the choke. You will remember that a choke is a device that opposes any change in current flowing through it. Therefore, the choke tries to keep the electrons leaving the positive side of C2 from flowing through it. Eventually this opposition offered by the choke is overcome, a magnetic field is built up in the choke, and some of the electrons can flow through the choke and capacitor C2 will be charged. However, the voltage to which C2 is charged will not be as high as the voltage to which C1 is charged because of the opposition of the choke.

When the ac input voltage drops below the voltage to which C1 is charged, the rectifier will no longer conduct and no current can flow through it. Now C1 and C2 and choke L1 must supply the current needed by the load. C2 does this by attempting to discharge. C1 also tries to discharge to charge C2 and to supply part of the current required by the load. At the same time there is a magnetic field built up in the choke L1 which does not collapse instantly but instead tries to keep current flowing in the direction it was flowing when the rectifier was conducting. It too helps to maintain the current flow through the load RL. Thus in this type of filter we have energy stored in three places: the capacitors C1 and C2 (as we did in the circuit in Fig. 13), and in the magnetic field of choke L1.

In the circuits shown in Fig. 15, the rectifier tube in Fig. 15A offers a certain amount of resistance to the flow of current through it. In addition, the rectifier tube has a cathode which must be heated by a heater. It takes some time for the cathode to come up to operating temperature; when a tube first starts conducting, the cathode is below normal operating temperature, and the tube offers a higher resistance to the flow of current through it than it does when the tube reaches its full operating temperature. Thus when the power supply is first turned on and the tube reaches a temperature at which the cathode begins to emit electrons, the tube offers considerable resistance to the flow of current through it. This limits the charging current through the tube that charges the input filter capacitor C1. In a matter of a few seconds capacitor C1 is charged and from then on the current that must flow through the tube is within the tube's capabilities.

In the circuit shown in Fig. 15B, however, the silicon rectifier does not have a cathode which must be heated -- as soon as the power is turned on the rectifier begins to conduct to charge the input capacitor C1. If you turn the equipment on at the peak of the ac cycle there will be a very high voltage immediately impressed across C1 and a very high current will flow through the rectifier. As a matter of fact, the current might be so high that it could burn out the rectifier. Even if the power is turned on when the ac voltage is at zero, a high current will flow through the diode to charge C1 as the voltage rises to a peak value with terminal A positive with respect to terminal B. If C1 is large enough, this could burn out the rectifier.

This problem of excessively high charging current can be overcome with the circuit shown in Fig. 17. Here a resistor is connected in series with the silicon rectifier to limit the current flow. In some equipment, this resistor is a fairly low resistance, fixed-value resistor. In other applications, the resistor may be a thermistor. You will remember that a thermistor is a resistor with a negative temperature coefficient. This means that the resistance of the thermistor decreases as its temperature increases.

With a thermistor for the resistor R1 shown in the circuit in Fig. 17, the thermistor will offer a fairly high resistance to current flow when the equipment is first turned on. This will limit the charging current that flows through the diode to charge C1 to a reasonable value. As the current flowing to the thermistor heats the thermistor and its resistance goes down, the charge across capacitor C1 will build up slowly so that the diode current never reaches an excessively high value. By the time the capacitor is fully charged, the thermistor temperature will have increased to a point where the resistance of the thermistor has dropped to a low value so that the thermistor has very little effect on the overall operation of the circuit.

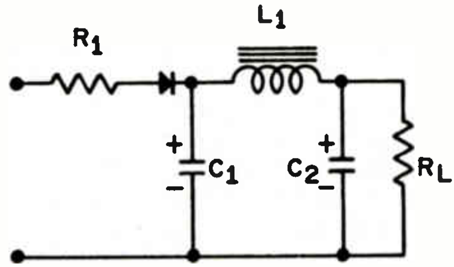


Fig. 17. Power supply with series-limiting resistor.

If you have to service a power supply of this type (where a resistor or a thermistor is used in the circuit in series with a silicon rectifier) and you find that the resistor or thermistor is opened, do not simply short the resistor or thermistor out of the circuit. If you do the chances are that the diode will burn out either the first time that you turn the equipment on or shortly thereafter.

The rectifier tubes used in low power equipment such as radio and television receivers are high vacuum rectifier tubes. However, in transmitters and in industrial applications where high voltages are involved you will often run into mercury-vapor rectifier tubes. Tubes of this type cannot be subjected to high peak currents through them without damaging the tube. A mercury-vapor rectifier tube in the circuit such as shown in Fig. 15A lasts only a short time. We can keep the peak current through the tube down to a safe value by using a somewhat different filter circuit known as a choke-input filter.

CHOKE-INPUT FILTERS

The filter circuits shown in Figs. 15 and 17 are called capacitor-input filters because the rectifier is connected directly to the input filter capacitor. A choke-input filter such as those frequently used with a mercury-vapor rectifier tube is shown in Fig. 18. Here, the rectifier tube is connected to a filter choke rather than a capacitor. Power supplies of this type will be found in radio and television transmitters and in industrial applications where comparatively high voltages are encountered.

It is easy to see how this type of filter works when we remember that the pulsating dc at the output of the rectifier tube is actually dc with ac superimposed on it. The choke offers little or no opposition to the flow of dc through it. On the other hand, the choke offers a high reactance to the flow of ac through it and at the same time the capacitor offers a low reactance to the ac across it. Thus the choke and the capacitor form a voltage divider network for the ac, so that most of the ac is dropped across the filter choke and very little of it appears across the load.

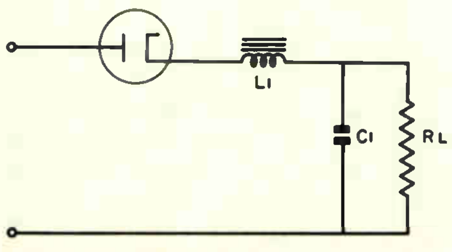


Fig. 18. A choke-input filter.

A more elaborate filter network is the two-section filter shown in Fig. 19. Again, it is a choke-input filter because the first element in the filter network is a choke. This type of filter network is frequently used in power supplies of radio and TV transmitters and of industrial electronic equipment where mercury-vapor rectifier tubes are used.

The choke-input filter has several advantages over the capacitor-input filter, even though the voltage obtained at the output of a choke-input filter is not quite as high as it is at the output of equivalent capacitor-input filters. In other words, if you feed the same pulsating dc into a choke-input filter you will not obtain as high an output voltage for a given load as you can with a capacitor-input filter. However, this type of filter has better voltage regulation than a capacitor-input filter. The voltage regulation of a power supply is the ratio of the full-load voltage to the no-load voltage. With a choke-input filter there is not as great a variation between the no-load and the full-load voltages as there is in the capacitor-input filter.

Another big advantage of the choke-input filter is that the peak current passed by the rectifier tube is held to a reasonable value. In a choke-input filter, the choke offers considerable reactance to any change in current flow through it. Thus, when the rectifier tube tries to conduct current heavily to charge C1 in Fig. 19, the input-filter choke L1 offers considerable reactance or opposition to the change in current flow through it. It tends to smooth

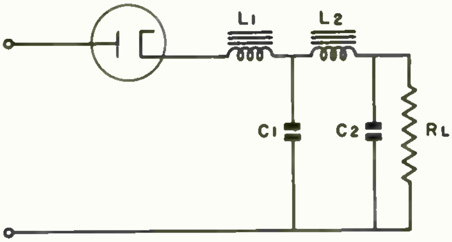


Fig. 19. A two-section choke-input filter.

out the pulses of current through the rectifier tube. The current pulse flowing through the rectifier tube flows for a slightly longer time than it would flow with an equivalent capacitor-input type of filter, and the peak amplitude of the current flowing through the rectifier tube is not as high as it would be with the capacitor input filter. This is a big advantage in a power supply using mercury-vapor rectifier tubes, because they can easily be destroyed by excessively high current pulses through them.

The filter network shown in Fig. 19 is quite effective in eliminating hum. Consider what would happen if the output of the rectifier had an ac voltage of 100 volts superimposed on the dc. If the two sections of the filter network are designed so that each choke has a reactance about 10 times as high as the reactance of the capacitors, each section will have approximately a 10-to-1 ripple voltage division; thus if there are 100 volts ac at the output of the rectifier, L1 and C1 will divide this voltage so that there will be only 10 volts appearing across C1. Now L2 and C2 act as a voltage divider network and divide this 10 volts further, so that the voltage across C2 would be only 1 volt. Thus, the two-section filter has reduced the ac hum or ripple voltage at the rectifier output from 100 volts to 1 volt.

The input filter choke, which is L1 in Fig. 19, is often a swinging choke. A swinging choke is designed so that it saturates rather easily and thus its inductance will vary appreciably as the current through the choke changes. When the current through the choke becomes high, the inductance and hence the inductive reactance of the choke decrease, but when the current through the choke is low, the inductance and hence the inductive reactance increase. Thus we have in effect a variable reactance between the rectifier tube and the input-filter capacitor; this variable reactance helps to improve voltage regulation at the power supply output. This type of choke is particularly useful in circuits where the load current goes through wide variations. If the load current goes down, the reactance of the choke increases. The increased reactance limits the charging action of the rectifier tube and keeps the output voltage from rising appreciably. On the other hand, if the load current increases, the reactance of the choke decreases, allowing the rectifier to charge C1 to a higher value so the capacitor can supply the increased current demand.

FACTORS AFFECTING THE OUTPUT VOLTAGE

In any filter network containing a filter choke or a filter resistor, the dc current flowing through the load must also flow through the filter choke or filter resistor. Thus, there will be a voltage drop across this choke or resistor; the exact value of the voltage drop will depend upon the dc resistance and the current flowing. We already pointed out that if a 1000-ohm filter resistor is used in the circuit and the current that is flowing is 100 milliamperes, the voltage drop across the filter resistor will be 100 volts. On the other hand, if a filter choke having a dc resistance of 100 ohms is used, a current of 100 milliamperes will produce a voltage drop of only 10 volts across it.

If the current drawn by the load changes, the output voltage at the output of the filter network will change. You can see why this is so--the current must flow through the filter resistor or filter choke. If the current flowing through the choke or resistor changes, the voltage drop across it will change, and hence the voltage at the output of the power supply will have some tendency to change.

The output voltage is also affected by the size of the filter capacitors used. If the filter capacitors used are small, then the rectifier is unable to keep them completely charged; if the filter capacitors are large, once the rectifier gets them charged, they will stay charged to a voltage near the peak ac voltage being applied to the rectifier tube.

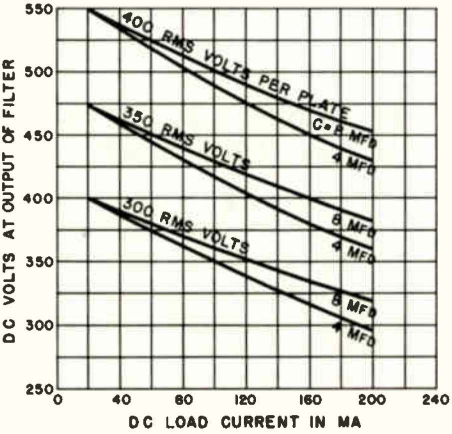


Fig. 20. How the size of the input capacitor of a filter affects the output voltage for different values of load current.

Of course, it is difficult to say whether a filter capacitor is large or small. Whether or not it is large for the particular circuit depends upon how much current is being drawn from the circuit. If the current drain is low, then a capacitor of 10 or 20-mfd will usually be sufficient to keep the power supply output voltage at or near the peak value of the ac voltage applied to the rectifier tube. However, if the current drain from the power supply is large, then the voltage across the power supply output will be considerably less than the peak ac value if the capacitors used are 10-mfd or 20-mfd capacitors.

To give you some idea of how the size of the input filter capacitor affects the output voltage, we have shown a graph in Fig. 20. The graph is for a full-wave rectifier. It shows how the output voltage varies with different load currents for both 4-mfd and 8-mfd capacitors. We have shown this for three separate input voltages. Notice that in each case when the current drain is low, the output voltage is substantially above the rms input voltage applied to the rectifier. Remember, of course, that the peak value of a 300-volt rms voltage is actually 1.4 times 300 volts.

Fig. 21 shows a comparison between a choke-input and capacitor-input type of filter. Notice that with a capacitor input the output voltage at low loads is substantially higher than it is for a choke input. However, as the load is increased, the output voltage from a capacitor-input type of filter drops rapidly. The output voltage from the choke-input type filter also drops as the load is increased, but it does not drop nearly as rapidly as it does with a capacitor-input type of filter.

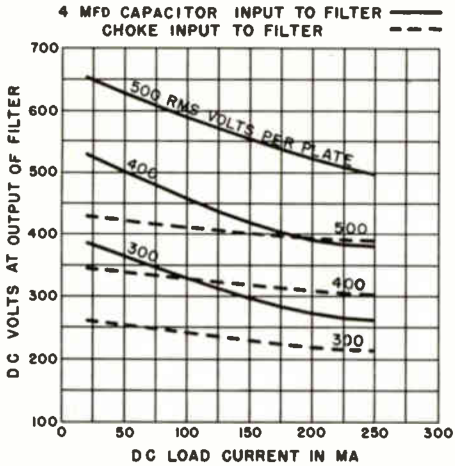


Fig. 21. How the output voltage of a filter network varies for different values of load currents and input voltages. The solid curves are for a capacitor-input filter; the dashed curves for a choke-input filter.

SUMMARY

In this section we have covered some of the more important types of filter networks you are likely to encounter in your career as an electronics technician. You have seen that these networks vary from comparatively simple filters consisting only of a capacitor up to networks containing two chokes and two filter capacitors.

Simpler types of filter networks can be used where the current drain is low and where the filtering does not have to be very good. The more elaborate filters are used in power supplies having a high current drain and in cases where good filtering is required to supply pure dc at the power supply output.

The circuits in this section of the lesson are shown with a half-wave rectifier. The same filter circuits are also used with full-wave rectifiers.

Study the circuits shown in this section. You should be familiar with each of these circuits. You might try drawing these circuits several times to get familiar with the circuit arrangement. Copy them from the book the first few times you try drawing them and then try to reproduce the circuit from memory. This will help you remember what the circuits look like, and if you can remember what they look like, the chances are that you'll be able to remember how they work.

SELF-TEST QUESTIONS

(i) To what value may the input filter capacitor charge in a simple capacitor filter such as shown in Fig. 11?

(j) In what type of application may a simple capacitor-type filter (such as shown in Fig. 11) be used?

(k) What advantage does an R-C pi-type filter (such as shown in Fig. 13) have over the simple capacitor-type filter shown in Fig. 11?

(l) What is the disadvantage of the R-C pi-type filter, and how can this disadvantage be overcome?

(m) What is the purpose of the resistor R1 in the power supply shown in Fig. 17?

(n) Why are choke-input filters used with mercury-vapor rectifier tubes?

(o) What advantage does a choke-input filter have over a capacitor-input filter?

(p) What is a swinging choke?

Typical Power Supplies

The two main sections of the power supply are the rectifier and the filter section. Now that we have discussed both these sections, let's examine some typical power supplies and see what they look like. First we'll look at some power supplies using tube-type rectifiers which might be found in equipment using vacuum tubes. Remember that in these circuits the rectifier operates in the same way as a selenium or silicon rectifier would operate. Insofar as the rectification action is concerned, it makes little difference whether a tube, a selenium rectifier, or a silicon rectifier is used. In the circuits where tubes are used, we will in some cases show how the tube heaters are connected. After looking at a few tube circuits we will look at some typical power supplies using silicon rectifiers and then at a more complex regulated supply.

UNIVERSAL AC-DC POWER SUPPLIES

The universal ac-dc power supply is so called because it can be operated from either an ac or a dc power line. When these power supplies were first used in radio receivers, manufacturers played this feature up, but actually this type of power supply is used to keep costs at a minimum.

The circuit of an ac-dc power supply for a 5-tube radio receiver is shown in Fig. 22. This supply not only supplies the dc voltages required by the plates and screens of the tubes in the receiver but also contains the heater supply for the heaters of the various tubes. First, let us look at the heater supply. Notice that the heaters of the various tubes are connected in series and that they are operated on ac. In this type of power supply you will always find that the heater of the rectifier tube, in this case a 35W4 tube, is connected directly to one side of the power line. The heater of this particular tube is tapped, and a pilot light is connected in parallel with one part of the heater. The pilot light will light when the set is turned on. The tube is designed to be operated with the pilot light connected in parallel with part of the heater, so if in servicing a receiver using a tube with a pilot light tap you should find that the pilot light is burned out, it is a good idea to replace it to keep the heater current in the rectifier tube within its rated value.

Next to the 35W4 rectifier tube in the heater circuit (or heater string, as it is usually called) you will always find the high-voltage power output tube. By a high-voltage tube we mean a tube that requires a high heater voltage. In the diagram we have shown, the tube is a 50B5 tube. As the 50 suggests, a heater voltage of 50 volts is required to operate the tube.

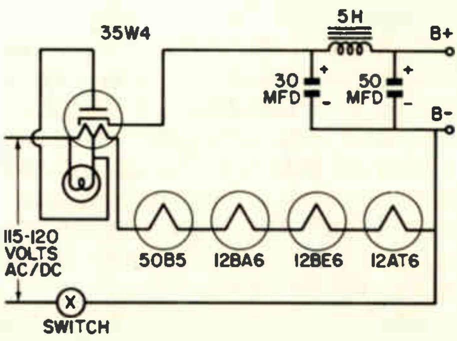


Fig. 22. A typical ac-dc power supply.

Next to the output tube you will usually find the i-f tube. In this diagram, the i-f tube is a 12BA6 tube. It is a tube with a 12.6-volt heater.

Next to the i-f tube you will find the converter tube. The 12BE6 tube is a converter tube designed so that the single tube performs two tasks: one part of the tube is the oscillator and the other part is the mixer.

The last tube in the heater string is always the first audio tube. In this circuit it is a 12AT6 tube. This tube is placed at the end of the heater string nearest B- to keep hum at a minimum, because hum picked up in the first audio tube will be amplified by the entire audio system and could be objectionable. The 12AT6 tube is a dual-diode-triode tube: it has two diodes and a triode inside the same glass envelope. As the 12 preceding the type designation suggests, the tube requires a heater voltage of about 12 volts--to be exact, 12.6 volts.

Now if we examine the rectifier circuit you will see that the plate is connected to a center tap on the rectifier heater. This means that the B supply current flowing through the plate of the tube must flow either through part of the rectifier heater or through the pilot light back to one side of the power line. The purpose of connecting the plate this way is to provide some protection in the event of a short in the receiver. If there is a short in the receiver, the rectifier tube will begin passing excessive current. If this current becomes too high, the pilot light and half the rectifier heater will burn out, opening the circuit and protecting the receiver and the house wiring.

The rest of the power supply is similar to the circuit you studied already. The filter network consists of a capacitor-input filter; the input filter capacitor is the 30-mfd capacitor, and the output filter capacitor is the 50-mfd capacitor. Usually these two capacitors are in a single container that has only three leads brought out from it. Since the negative leads are both connected to B-, a common negative lead and two separate positive leads are usually used. The capacitor is usually a tubular type of capacitor with a paper cover impregnated with wax. The capacitor is mounted on the receiver chassis by means of a mounting strap, or in a receiver with a printed circuit board, the leads are soldered directly to the circuit board. Often this is the only kind of mechanical mounting used.

Capacitors in this type of receiver are usually rated at 150 volts. The normal output voltage of this type of power supply under load is usually somewhere between 90 and 105 volts.

Not all universal ac-dc power supplies use a filter choke. In the circuit shown in Fig. 22, the plates and screens of all of the tubes are operated from the B+ output of this power supply. However, sometimes you will find a filter resistor used in the power supply in place of the filter choke. When this is done, the plate of the output tube is usually connected directly to the cathode of the rectifier tube so that the current drawn by the output tube plate will not flow through the filter resistor. This tube usually draws more current than all the other tubes in the receiver combined.

The output tube is normally a beam tube or a pentode tube, and the plate current depends very little on plate voltage. Therefore if there is some hum voltage applied to the plate of the tube, it usually does not cause any plate current variation and hence is not heard as hum in the output from the loudspeaker. The screen voltage for the output tube and the plate and screen voltages for the remaining tubes are obtained at the output of the power supply where the additional filtering obtained from the filter resistor and the output filter capacitor will reduce the ripple from the rectifier to a low value.

In some of the later table-model radio receivers only four tubes are used; a selenium rectifier or a silicon rectifier is used in place of a rectifier tube because either of these has a considerably longer life. In receivers of this type, tubes with higher heater voltage requirements are often used, so that the sum of the heater voltages required by the four tubes adds up to about 120 volts. If the heater voltage required by the tubes is less than the line voltage, a voltage-dropping resistor can be placed in series with the heaters to use up the leftover voltage. This will cause the total voltage drop across the series-dropping resistor and the tubes to be equal to the line voltage, thus allowing each of the tubes to have its required heater voltage.

Of course, in any heater string where tube heaters are connected in series, all the tubes must have the same heater current rating. Tubes designed for this type of service have what is called a controlled warm-up. This means that the tube heaters are made so that they all reach operating temperature at the same time. This prevents one tube from warming up too quickly and from having too high a voltage across its heater.

The voltage across the tube heater will depend upon its resistance, which changes with temperature. As long as all the tubes warm up at the same rate, their resistances change at the same rate. In this way, a high voltage across any of the tubes is avoided.

Before leaving our study of this type of power supply, it is worthwhile to consider what will happen if a tube such as the 12BA6 tube in Fig. 22 develops a cathode-to-heater shortage. The cathode will usually be connected either to B- directly or to B- through a low-value resistor; this will effectively short out the heaters of the 12BE6 and 12AT6 tubes and, as a result, these tubes will not light.

If you're called upon to service a receiver of this type and you find that one or more of the tubes is not heating, look for a cathode-to-heater short in either the tube which doesn't light that is highest up on the string, or the tube preceding it. If none of the tubes lights this is an indication, since the tubes are connected in series, that the series string is open -- chances are that the heater of one of the tubes has burned out.

A TYPICAL FULL-WAVE POWER SUPPLY

Fig. 23 illustrates a typical power supply using a power transformer and full-wave rectifier. This type is found in many radio and TV receivers and transmitting equipment, as well as in the equipment used in industrial electronics.

Notice that the power transformer has a primary winding to which the 115-volt ac power line is connected; a high-voltage secondary with its center tap grounded, and the two end leads connected to the plates of the rectifier tube; and the two filament

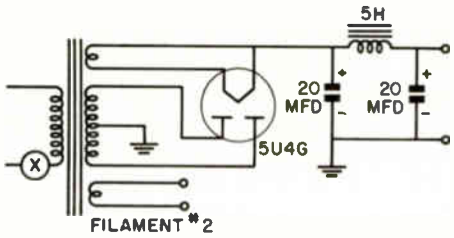


Fig. 23. A typical power supply using a transformer and full-wave rectifier.

windings. One filament winding is used to supply the heater voltage required by the rectifier tube, and one filament winding is used to supply the heater voltage for all the other tubes in the set.

In this power supply we have shown a 5U4G rectifier tube, which requires a heater voltage of 5 volts and a heater current of 3 amps. Thus the rectifier filament winding on the transformer must be capable of supplying 5 volts at a current of 3 amperes.

The winding marked filament number two is used to supply the heater voltage required by all the other tubes in the set. You will notice that this winding is called the filament winding, not the heater winding. This is a carry-over from the old days of radio when most of the tubes were filament-type tubes and few had a separate cathode and heater.

In an electronic device using this type of power supply, the heaters of the tubes (with the exception of the rectifier tube) are connected in parallel. Since the tubes are connected in parallel, you will find that all of the tubes are designed to operate from the same heater voltage. In most cases equipment using this type of power supply has tubes with heater voltage ratings of 6.3 volts. The filament winding must be capable of supplying the heater current required by all of the tubes. The tubes may require different heater currents since they are connected in parallel. You can determine the total current the winding must supply by looking up the heater current required by the individual tubes in a tube manual and adding these figures together.

Again, the filter network used in this power supply is a capacitor-input type. Notice that the capacitors have a lower capacity than those used in the circuit shown in Fig. 22. It is not as necessary to use large capacitors in a full-wave rectifier type of power supply as it is in a half-wave rectifier to obtain the same amount of filtering. Of course, in some power supplies in which it is essential to keep the hum voltage very low, you will find larger filter capacitors. You may also find a two-section filter using an additional choke and a third filter capacitor.

In a power supply such as the one shown in Fig. 23, the two capacitors will probably be mounted in a single container. If a cardboard tubular type is used, one common negative lead and two separate positive leads will be brought out of the container. The same color leads will probably be used for the two positive sections since they both have the same capacity. In some pieces of equipment a metal can-type capacitor might be used -- with this type the can is usually the negative terminal. Mounting the capacitor on the metal chassis automatically makes the connection between the negative terminal and the chassis. The positive leads are brought out of two separate terminals.

DUAL-VOLTAGE POWER SUPPLIES

A diagram of a typical dual-voltage power supply is shown in Fig. 24. This diagram is actually quite similar to the diagram of circuits used in a modern color-TV receiver. Notice that a voltage that is negative with respect to ground is developed by the diode D1 and its associated circuitry. Diodes D2 and D3 are used in a half-wave voltage doubler circuit.

In this circuit when terminal A is negative with respect to terminal B, current flows from A through R1 and D1 to charge capacitor C1. A simple pi-type R-C filter is used in this section of the power supply because the current requirements are low and there will be very little voltage drop across R2. At the same time with the two capacitors, small value capacitors can be used and adequate filtering obtained.

When terminal A is negative, current also flows through the thermistor R3 through R4 and into the negative plate of capacitor C3. Electrons flow out of the positive plate through the diode D2 back to terminal B of the power line. When terminal A of the power line is positive and terminal B is negative the voltage will be placed in series with the voltage built up across C3, so that capacitor C4 is charged to a value approaching twice the peak line voltage through diode D3. This provides an output voltage which is positive with respect to ground and approximately equal to twice the ac line voltage.

The thermistor R3 is put in the circuit to prevent high current surges through the silicon diodes D2 and D3 when the equipment is first turned on. R4 is used in the circuit to provide further protection. If the equipment is turned on and operating for some time, the resistance of the thermistor will drop to a low value. If the equipment is turned off for a

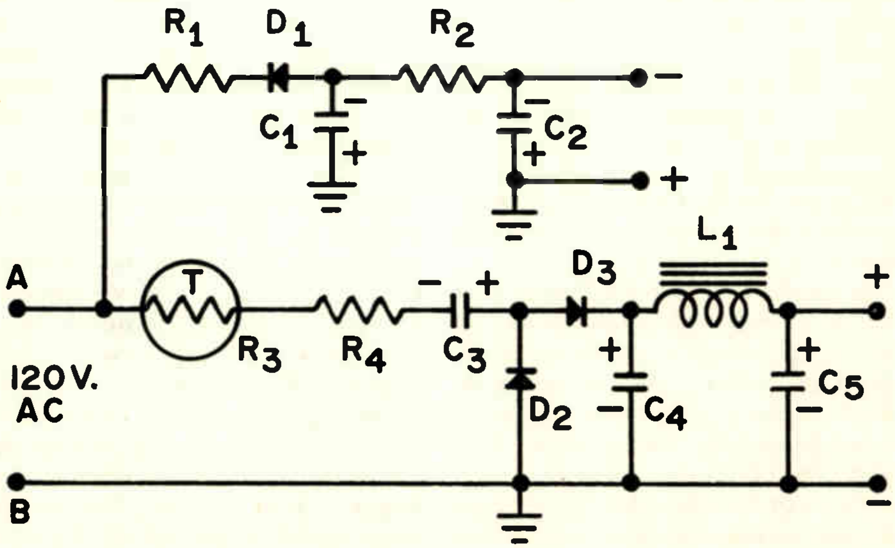


Fig. 24. A typical dual-voltage power supply.

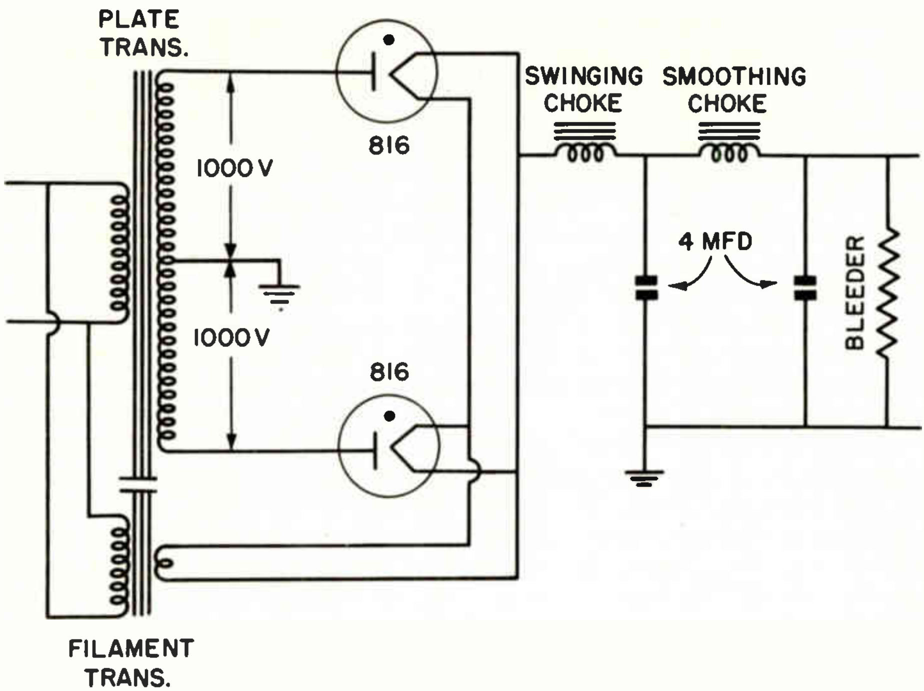


Fig. 25. A typical transmitter power supply.

few seconds and then turned back on, the charging current through the diodes could be quite high and damage them. R4 is put in the circuit; its value does not change and it limits the current through the diodes to prevent their damage under this circumstance.

The power supply shown in Fig. 24 is typical of the power supply you are likely to run into in both monochrome and color television receivers. A voltage-divider network may be used across the output of either supply to provide different value voltages. This power supply is comparatively inexpensive; it eliminates the need of a power transformer and with modern silicon diodes is comparatively trouble-free.

HIGH-VOLTAGE POWER SUPPLIES

Another power supply is shown in Fig. 25. This is the type of power supply you will find in transmitting equipment or in other equipment where high operating voltages are required. Notice that in this power supply, full-wave rectification is used. Also notice that it has two separate rectifier tubes. Separate rectifier tubes rather than a single tube with two plates are used in high-voltage supplies because the voltages are so high that they would simply arc across inside the tube. The type 816 tubes shown in this supply are mercury-vapor tubes designed for use in power supplies where the operating voltages are not too high and where the current requirements are not too great. They are often used in power supplies where the ac input voltage to the rectifier is between 1000 and 2000 volts, and the current drain does not exceed 250 ma. As far as the transmitting-type power supplies are concerned, an output voltage of 1000 volts across each half of the secondary of the high-voltage transformer is not considered high.

The high-voltage transformer found in this type of power supply is frequently called a plate transformer because it is used to supply the high voltage required to operate the plates of the various tubes in the transmitter. The rectifier tubes are operated from a separate filament transformer. In a power supply using type 816 tubes, the filament voltage required by these tubes is 2.5 volts and the current required is 2 amps. Therefore, this filament transformer must be capable of supplying a total of 4 amps at a voltage of 2.5 volts. The filament transformer must have good insulation between the secondary winding, the transformer core, and the primary winding; otherwise the high voltage will arc through the insulation either to the primary winding or to the transformer core.

At the input of the power supply filter network is a swinging choke. This is common practice in transmitter power supplies because it improves the voltage regulation, and also because it affords additional protection for the mercury-vapor rectifier tubes.

The second choke in the power supply is called a smoothing choke. This is the same type of choke shown in the power supplies in Figs. 22 and 23. It is called a smoothing choke in transmitting equipment because its primary purpose is to smooth out the ripple and also to distinguish it from the swinging choke used at the input of the filter network. A smoothing choke is designed to keep its inductance as nearly constant as possible, whereas a swinging choke is designed so that its inductance will vary as the current through it varies.

Notice that the filter capacitors used in this power supply are 4-mfd capacitors. Each of these capacitors is usually mounted in its own separate container. Occasionally you will find two small high-voltage capacitors in the same container, but this is not common practice. Also notice that the capacitors have a much smaller capacity than those in the two power supplies we discussed previously. These capacitors are oil-filled capacitors and it is quite costly to make this type of capacitor with a large capacity. On the other hand, capacitors used in circuits like those in Fig. 22, Fig. 23, and Fig. 24 are electrolytic capacitors and very large capacities can be obtained at a very moderate cost.

Effective filtering is obtained with the smaller-size capacitors in this supply, because two filter chokes are used. The inductance and hence the inductive reactance of these chokes are usually somewhat higher than that of filter chokes used in lower voltage equipment. The choice of using either large chokes or large filter capacitors is simply one of cost. At low voltages it is more economical to use large capacitors and low-inductance chokes, but at high voltages it is more economical to use high-inductance chokes and low capacities. The net result is the same as far as the filtering action is concerned. Another component that you will find in transmitting-type power supplies is a bleeder. A bleeder is a resistor connected across the power supply output. The bleeder serves two purposes: it improves the voltage regulation and is also used for safety.

A bleeder resistor connected across the power supply keeps the minimum current at a reasonable value. If the current drawn by the load connected to this power supply were to drop to zero, the two filter capacitors would be charged up to a value equal to the peak voltage across half of the secondary of the plate transformer. If the transformer had an rms voltage of 1000 volts across each half of the secondary, this would mean that the capacitors would charge up to a voltage of about 1400 volts. This may be high enough to destroy the capacitors. Also, the chokes have a definite maximum voltage that can be applied to them. If the voltage goes too high, the insulation between the choke winding and the core may break down. This will destroy the chokes. If either of the chokes or capacitors shorts, the rectifier tubes will pass such a high current that they may be ruined. There is also the danger of burning out the plate transformer. Furthermore, if the voltage reaches too high a value, the rectifier tubes may arc over internally. A bleeder connected across the power supply output can eliminate this danger. With the bleeder across the output, if the equipment current drops to or almost to zero, the bleeder current will continue to flow. If the output voltage starts to rise, the bleeder current will increase because the current flowing through any resistor increases if the voltage across it increases. The bleeder current will keep the voltage from climbing to an unsafe value.

Another important reason for using the bleeder is that an oil-filled capacitor such as those found in this type of power supply can hold a charge for a long time. If the two 4-mfd capacitors used in the supply were charged up to a voltage of 1000 volts or more and a technician servicing the equipment accidentally touched one of these capacitors, he could receive a very dangerous shock. Under certain conditions it could be fatal. This danger can be greatly reduced by connecting a bleeder across the power supply so that when the equipment is turned off the capacitors are discharged through the bleeder.

You may have occasion to work on high-voltage power supplies at some future date. Remember that a bleeder is connected across the power supply for safety as well as to improve the voltage regulation. Therefore if the bleeder in a power supply burns out, it should be replaced. However, never rely on a bleeder to discharge high-voltage filter capacitors. If you have to work on a high-voltage power supply, your first step should be to remove all voltages from the supply. To do this, turn the power supply off; if there are fuses in it, remove the fuses so that no one can accidentally turn it on; disconnect it completely from the source if you can. Often it is not possible or convenient to completely disconnect the equipment from the voltage source but if it is shut off and any fuse in the circuit is removed, it should be safe. Next, before you start to work on the supply, discharge all filter capacitors in the power supply. The capacitors should be discharged with a heavy metal rod that has a good insulated handle so that you will not come in contact with the metal rod. Use the metal rod to short together the terminals of the capacitor to discharge it. Touch the grounded terminal of the capacitor first and slide the rod over to touch the other terminal. Do this several times to be sure the charge is completely removed. After the capacitors have been discharged, the power supply should be safe to work on.

Keep this point in mind: high voltage capacitors, or for that matter any large capacitors, should be discharged before you start to work on a piece of equipment. Many technicians fail to do this. There are some technicians who can tell about the terrific shock they received when they failed to discharge a filter capacitor. There are others that did not survive the experience to tell about it.

A TRANSISTOR-REGULATED POWER SUPPLY

A regulated power supply employing transistor voltage regulators is shown in Fig. 26. This power supply is used in a TV receiver that is designed for operation from the power line and also from a 12-volt dc source. When the power plug is plugged into a 120-volt line and the switch is turned on, the receiver will operate from the power line. When the power plug is disconnected and switch S1 is closed, the receiver can be operated from a 12-volt battery.

The operation of the power supply from the power line is comparatively simple. Two diodes, D1 and D2, are used in a full-wave rectifier circuit. When the transformer T1, which is a step-down transformer, has a polarity such that the end of the secondary connected to D1 is negative, current will flow through the diode D1 to ground and into the negative plate of C3. Electrons flow out of the positive plate of C3 to the center tap of the power transformer which is positive with respect to the end connected to D1. During the next half-cycle, when the end of the secondary connected to D2 is negative, current will flow through D2 to ground, into the negative plate of C3,

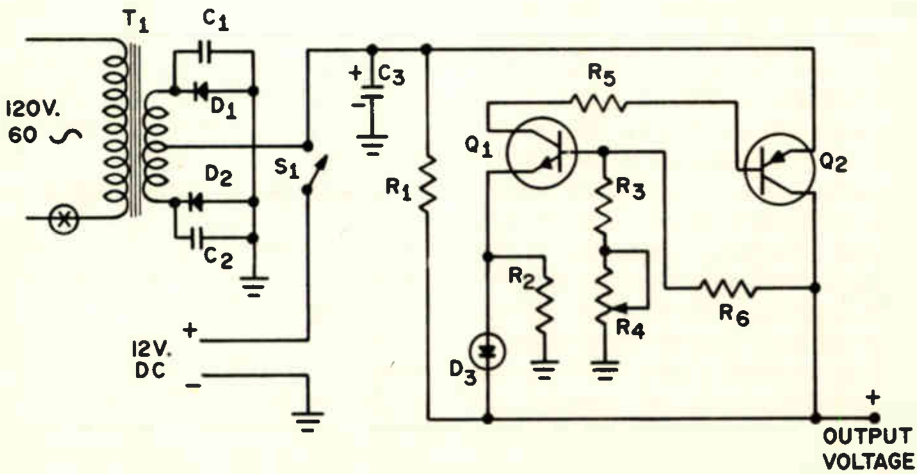


Fig. 26. A transistor-regulated power supply.

out of the positive plate of C3 and back to the center tap of the secondary winding on the power transformer.

The remainder of the components used for the power supply are used for the purpose of regulating the voltage. In other words, the power supply voltage is maintained constant at approximately 12 volts regardless of the load drawn from the supply. The transistor Q2 is a PNP transistor that is used as a series voltage regulator. Notice that the emitter of this transistor connects directly to the positive side of C3. You can consider this transistor as working more or less as a variable resistor: if the output voltage tends to rise, the resistance increases and if the voltage tends to fall, the resistance decreases.

The effective resistance of Q2 is varied by varying the forward bias across the emitter-base junction. Notice the zener diode D3. This diode is connected in series with R2. The zener has a constant voltage of 6.3 volts across it. Therefore, the voltage drop across R2 will be equal to the output voltage minus 6.3 volts. This is the emitter voltage applied to Q1. The base voltage is determined by the voltage division occurring between R4, R3 and R6. R4 is adjustable so that the output voltage can be adjusted to 12 volts. Under these circumstances a certain current will flow through Q1 and through R5 and this will set the forward bias on Q2. If the output voltage tends to rise, the base voltage on Q1 will rise but by an amount less than the emitter voltage. The divider network consisting of R4, R3 and R6 will prevent the base from rising the full amount of the output voltage rise. On the other hand, the voltage across the zener D3 remains constant so that the voltage across R2 will reflect the entire output voltage rise. This will reduce the forward bias on Q1 which, in turn, will reduce the emitter-collector current. The reduction in the emitter-collector current will reduce the voltage drop across R5 which, in turn, will reduce the forward bias on Q2; this has the effect of increasing its resistance. The increased resistance tends to keep the output voltage from increasing.

If the output voltage decreases, the opposite happens. The base voltage on Q1 falls, as does the emitter voltage. However, the emitter voltage falls more than the base voltage, so the forward bias is increased. This increases the emitter-to-collector current through Q1 which increases the forward bias on Q2. This has the effect of reducing the resistance on Q2 and tends to keep the output voltage from falling.

This type of power supply is one of the more complex power supplies that you are likely to encounter in electronic equipment. The voltage regulation is required in order to keep the voltage reasonably constant on the various transistors used on the TV receiver. In most cases, such precise voltage regulation is not required in entertainment-type equipment.

VOLTAGE DIVIDERS

We mentioned previously that more than one operating voltage is sometimes needed in the various stages of a piece of electronic equipment. Rather than use a separate supply for each voltage needed, the usual procedure is to use a single supply designed to give the highest voltage needed, and then obtain the lower voltages required by means of a voltage divider connected across the power supply output. A typical voltage divider is shown in Fig. 27.

In this voltage divider, R1 and R2 are voltage-dropping resistors; they drop the voltage from 300 volts to the required voltages of 200 and 100 volts. R3 is a bleeder used to stabilize the voltages at points B and C. With this type of network, terminal D is the ground or common terminal. Between D and C there is a voltage of 100 volts; terminal C is positive with respect to terminal D. Between D and B there is a voltage of 200 volts and terminal B is positive with respect to terminal D. Finally, between terminals D and A there is the full power supply output voltage of 300 volts, and of course terminal A is positive with respect to terminal D.

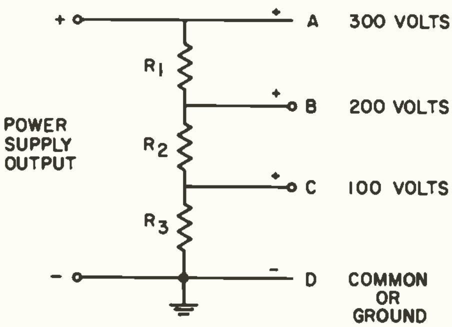


Fig. 27. A voltage-divider network.

The current flowing through R3 is called the bleeder current. It remains fairly constant and is determined primarily by the sizes of R1, R2 and R3. Usually, the size of R3 is chosen so that the bleeder current will be at least as great as the current drawn by the stages connected to terminals C and B. Choosing a value of R3 that will result in a reasonable bleeder current helps maintain good voltage regulation at terminals C and B.

The current flowing through R2 will be made up of the bleeder current plus the current drawn by the stages connected to terminal C. If this current varies, the voltage drop across R2 and hence the voltage at terminal C will vary. However, the bleeder current will remain essentially constant so that if a sizable percentage of the current flowing through R2 is bleeder current, variations in the current drawn by the stages connected to terminal C do not cause too much variation in the voltage drop across R2.

The current flowing through R1 is made up of the bleeder current plus the current drawn by the stages connected to terminals C and B. Again if the bleeder current through R1 represents a sizable part of the total current flow through R1, variations in the current drawn by the stages connected to terminals C and B do not cause too great a variation in the voltage drop across R1 so the voltage at terminal B will remain reasonably constant.

Bleeders are not used in modern midget radio receivers, but you will find them in many of the older sets. They are frequently used in TV receivers, in the low-voltage power supplies in transmitting equipment, and in industrial electronic equipment.

Sometimes one section of a tapped resistor will burn out. Often you can repair the equipment simply by connecting a resistor having the correct resistance and a suitable wattage rating across the defective section. Of course, if separate resistors are used in the voltage divider you can simply replace any defective one. If you do shunt a burned out section of a tapped resistor in a radio receiver and find the equipment is noisy after you have made this repair, the defective section may be making contact intermittently and creating the noise. Of course in this case you must replace the entire unit either with separate resistors connected in series or with a tapped resistor like the original one.

VIBRATOR-TYPE SUPPLIES

The radios installed in automobiles for years used a power supply known as a vibrator type of power supply. A schematic diagram of this type of power supply is shown in Fig. 28. The heart of this type of power supply is the vibrator, which is used to change the dc from the automobile storage battery to a pulsating current in the primary winding of the power transformer.

The vibrator consists of an electromagnet L, and a reed (R-K) placed between two sets of contacts. In the circuit shown in Fig. 28, when the switch is turned to the ON position, current will flow from the negative terminal of the battery through the switch and through the reed towards terminal M. Here it will flow from the reed to contact M, to coil L, through coil L back to the positive side of the battery. The current flowing through the coil creates a magnetic field. This magnetic field attracts the end of the reed K, pulling the reed over toward L and contact N. When the reed makes contact with terminal N, current flows through the upper half of the transformer primary winding. It flows from the top of the winding to the center tap, building up a magnetic field.

At the same instant that the reed is making its contact with terminal N, it will break its contact with terminal M so that the electromagnet will no longer be energized and the field about it will collapse. The reed is made of a spring type material so that it springs back until it makes contact with both terminal M and terminal O. At the instant contact is made with terminal O, current flows through the lower half of the primary winding of the transformer, flowing from the bottom of the winding towards the center tap. The current is flowing through the primary winding in the opposite direction to the direction in which it was flowing through the upper half of the trans-

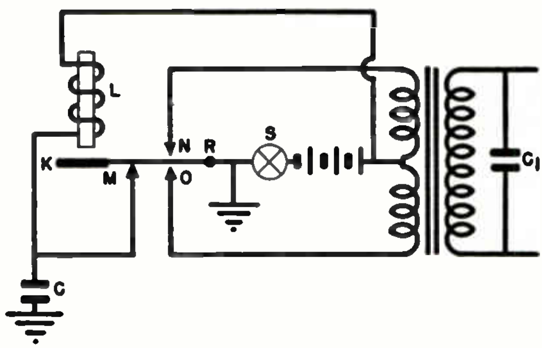


Fig. 28. A vibrator circuit.

former winding. Thus we have a field built up in one direction and then in the opposite direction. At the same time the fact that the reed makes contact with terminal M will once again complete the circuit through the electromagnet so the reed will swing over to the magnet again, making contact with terminal N. As you can see this action causes the reed to vibrate back and forth between terminals N and O. Thus we have a field built up in the primary first in one direction and then in the opposite direction. Building up this field, collapsing it, and then building up a reverse field and collapsing it, means that we have a continually changing magnetic field cutting the secondary of the transformer. By putting a large enough number of turns on the secondary, we can obtain whatever voltage we may require for the operation of the receiver.

A complete vibrator-type power supply is shown in Fig. 29. Notice that the secondary of the vibrator transformer is center tapped and that a full-wave rectifier is used. The capacitor C_2 is called a buffer capacitor. This is a high voltage paper capacitor that is used to keep sharp noise pulses out of the power supply. The size is usually quite critical and if it is necessary to replace the buffer capacitor in a receiver using this type of power supply, you should use a capacitor having the same capacity as the original.

In a vibrator-type power supply there is considerable sparking as the reed vibrates back and forth. This sets up a radio-frequency type of interference which could get through the rectifier and cause considerable interference in the receiver. In the power supply shown in Fig. 29, the choke L and the capacitor C are called hash suppressors. This rf interference or noise is called hash; the choke and the capacitor are put in the power supply in order to keep as much as possible of this hash or noise out of the power supply output. Capacitor C acts like a short circuit to these radio frequency pulses, and choke L acts like a very high impedance to them. Thus L and C form a voltage-divider network, with most of the voltage appearing across the high impedance L and little or no voltage across the low impedance C.

Vibrator-type power supplies were used in almost all automobile receivers in automobiles using 6-volt ignition systems. However, in newer cars, a 12-volt ignition system is used. The first receivers made for these cars also used vibrator type supplies, but tube manufacturers designed special tubes that will operate with plate and screen voltages as low as 12 volts. These 12-volt tubes were used in automobile receivers for a few years, but they too were replaced by transistors. Since transistors operate from low voltages, the vibrator type of power supply is no longer needed. These supplies were not only costly, but in addition they were one of the most troublesome sections of the automobile receiver.

SUMMARY

The power supplies we have shown in this section of this lesson are typical of the various types of power supplies you are likely to encounter as an electronics technician. You will find many variations of these circuits, but these are the basic circuits. Spend some time studying

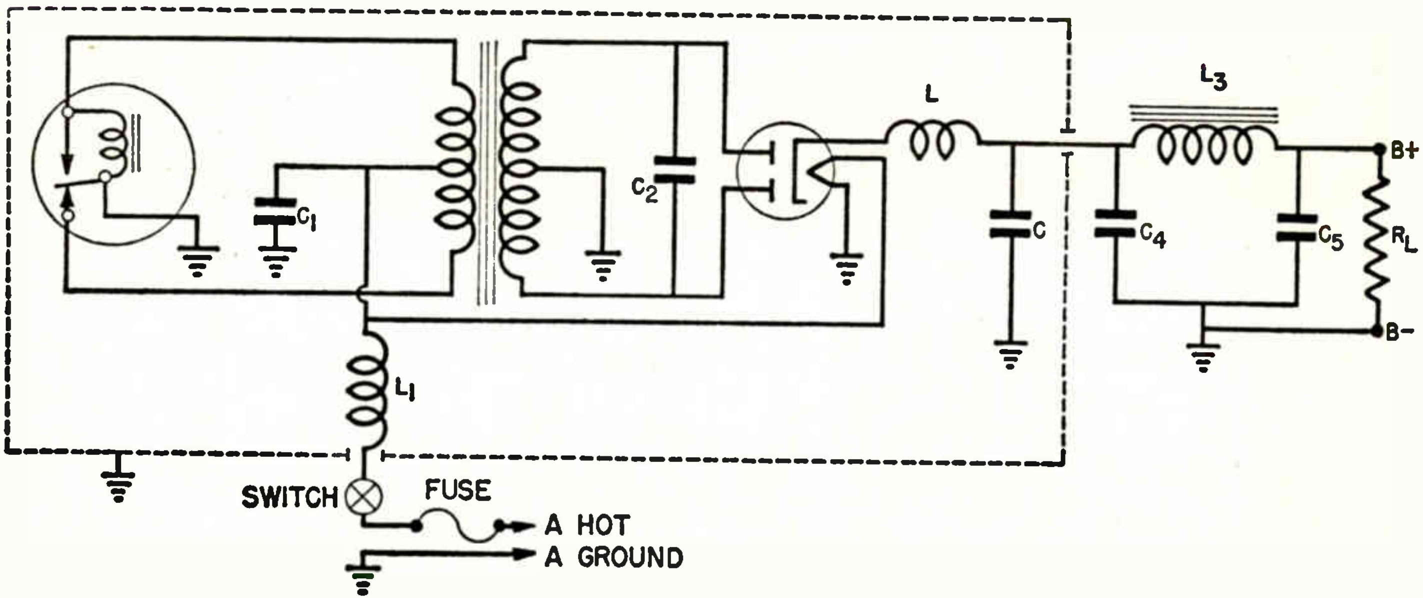


Fig. 29. A complete vibrator power supply.

these circuits so you will know what they look like.

In servicing these power supplies, keep in mind that the values of the various components used are generally not extremely critical. Manufacturers usually use as small a filter capacitor as they can in a circuit. It is more economical to use a small capacitor than to use a large one. Therefore, if you have an ac-dc receiver that uses a 20-40 mfd, 150-volt capacitor and you find it necessary to replace the filter capacitor, there is no reason why you could not use a 30-50 mfd, 150-volt capacitor in its place. Some variation of part values can be made without affecting the performance of the equipment.

It is also possible to use capacitors having a higher voltage rating than those used originally. This will simply provide an additional margin of safety. You may find that you cannot do this in large transmitter power supplies because there isn't room to mount a capacitor with a higher voltage rating, but usually in radio, TV, and most small pieces of industrial electronic equipment there is enough room to put in a capacitor of slightly larger physical size than the original.

SELF-TEST QUESTIONS

(q) Draw a schematic diagram of a typical ac-dc power supply found in a five-tube radio.

(r) In the circuit shown in Fig. 22, why is the 12AT6 tube placed at the B- end of the string?

(s) If in an ac-dc receiver in which the heaters of the tubes are connected in series and none of the tubes lights what would you suspect the cause of the trouble to be?

(t) Why are the filter capacitors used in the power supply shown in Fig. 23 smaller in capacity than the filter capacitors used in the power supply shown in Fig. 22?

(u) What is the purpose of R3 in Fig. 24?

(v) What is the purpose of R4 in Fig. 24?

(w) What type of voltage-doubler circuit is used in the power supply shown in Fig. 24?

(x) What purpose does the transistor Q2 serve in the power supply shown in Fig. 26?

(y) What is the purpose of the diode D3 in Fig. 26?

ANSWERS TO SELF-TEST QUESTIONS

(a) There will be 60 current pulses per second through the load.

(b) The disadvantage of the half-wave rectifier circuit is that current flows through the rectifier during one half-cycle and not during the other half-cycle. As a result, the output is somewhat difficult to filter and smooth out to pure dc.

(c) The diode D1 in the circuit shown in Fig. 6 is used to charge the capacitor C1 during one half of each cycle. Capacitor C1 is charged so that during the next half-cycle the voltage across it will be in series with the line voltage. This will place a voltage equal to twice the line voltage across the load and diode D2. Since diode D2 has a very low resistance when it is conducting, the voltage across the load is twice what it would be without the combination of C1 and D1 in the circuit: hence, the circuit is called a voltage-doubler circuit.

(d) The circuit is called a full-wave rectifier circuit because a current pulse flows through the load during each half-cycle. In other words, if the rectifier circuit is operating from a 60-cycle power line there will be 120 current pulses through the load (one for each half-cycle).

(e) The disadvantage of this circuit is that the high voltage winding on the power transformer must be center-tapped. This means that the high-voltage winding on the transformer must have twice the number of turns required to get the desired output voltage across the load. This circuit requires a rather expensive power transformer.

(f) The advantage of the bridge rectifier circuit is that there is a saving in the power transformer cost over that of a transformer that has a tapped high-voltage secondary winding; also, the circuit is capable of good voltage regulation.

(g) The voltage-doubler circuit shown in Fig. 9 is a full-wave voltage doubler. That is, there will be 120 current pulses per second in the output of the voltage doubling capacitor network consisting of C1 and C2. The circuit shown in Fig. 6 is a half-wave voltage-doubler circuit and there will be only 60 current pulses through the load in this circuit. It will be somewhat easier to filter and smooth the output voltage in the circuit shown in Fig. 9 than it will be in the circuit shown in Fig. 6.

(h) A full-wave doubler circuit requires a less expensive power transformer for a given load voltage than the bridge rectifier circuit requires. Also, the voltage-doubler circuit requires only two rectifiers whereas the bridge-rectifier circuit requires four rectifiers. The disadvantage of the full-wave voltage-doubler circuit is that it does not have as good voltage regulation as the bridge-rectifier circuit.

(i) The capacitor in a simple filter circuit such as shown in Fig. 11 may charge up to a value equal to the peak value of the ac input voltage. In the case of a power supply operating from a 120-volt line this is equal to approximately 1.4 times 120 volts.

(j) A simple capacitor-type filter may be used in applications where the current drain is low. With a low current drain the capacitor discharges very little between cycles so that the voltage across the capacitor, and hence the voltage across the load, remains essentially constant.

(k) The R-C pi-type filter is capable of better hum elimination than a simple capacitor-type filter. This type of filter is particularly desirable where the current drain is high enough to discharge the capacitor appreciably between charging cycles in a simple capacitor-type filter.

(l) The disadvantage of the R-C pi-type filter is that there is considerable voltage drop across the filter resistor. This problem can be overcome by using a filter choke such as in the L-C type filter shown in Fig. 15. A filter choke will offer a high opposition to any ac and thus effectively reduce the ac, while at the same time offering a low resistance to the passage of dc through it.

(m) R1 in the power supply shown in Fig. 17 is used to limit the current through the silicon rectifier when the power supply is first turned on. Without this resistance in the circuit, the charging current through the diode to charge C1 may be so high that the rectifier may be destroyed.

(n) To limit the peak current through the tubes. A mercury vapor rectifier tube is easily damaged by a high peak current. The peak current through the rectifier tube is much lower with a choke-input filter than it is with a capacitor-input filter.

(o) A choke-input filter will provide better regulation than a capacitor-input filter. This means that the voltage across the load will vary less with widely varying currents when the filter is a choke-input filter than it will when the filter is a capacitor-input filter.

(p) A swinging choke is a choke whose inductance changes as the current changes. As the current builds up the choke tends to saturate so that its inductance goes down. This tends to reduce the reactance of the choke and hence helps provide better voltage regulation.

(q) See Fig. 22. If you cannot draw this diagram from memory, copy it from the book. Simply drawing the diagram will help you to become familiar with the circuit and remember it in the future.

(r) The 12AT6 tube is the first audio stage. It is placed at the B- end of the heater string in order to keep hum pick-up in the tube as low as possible. Any hum picked up by this tube will be amplified by the entire audio system.

(s) The chances are that the heater of one of the tubes is open.

(t) The power supply shown in Fig. 23 uses a full-wave rectifier. Therefore there will be 120 pulses per second to charge the filter capacitors. The power supply shown in Fig. 22 is a half-wave power supply and there will be only 60 pulses per second to charge the filter capacitor; therefore, larger capacitors are needed to eliminate hum.

(u) R3 in Fig. 24 is a thermistor. A thermistor has a high cold resistance, but the resistance decreases as the thermistor heats up. The thermistor is used in this power supply to protect the diode rectifiers from high current surges when the power supply is first turned on.

(v) R4 is a fixed resistor that is used to protect the diodes in the event the equipment is turned off and then turned back on almost immediately. Under these conditions the resistance of the thermistor will be too low to provide the required protection for the rectifiers and hence R4, along with the hot resistance of the thermistor, limits the current through the diode rectifiers to a safe value.

(w) A half-wave voltage doubler.

(x) Q2 is in series with the B supply voltage. It operates essentially as a variable resistor and is used to regulate the power supply output voltage and keep it at essentially a constant value.

(y) The diode D3 is the zener diode. It provides a reference voltage so that the voltage variations on the emitter of Q1 will be greater than the voltage variations on the base. Thus, changes in output voltage affect the forward bias of the transistor and hence the conduction through it.

Lesson Questions

Be sure to number your Answer Sheet B201.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What advantage does a silicon rectifier have over a vacuum-tube rectifier?

2. Draw the schematic symbol for a silicon rectifier, and indicate by an arrow the direction in which the current will flow through it.

3. How many pulses per second will you get at the output of a full-wave rectifier that is operated from a 60-cycle power line?

4. What is the purpose of C1 and D1 in the circuit shown in Fig. 6?

5. (a) In the circuit shown in Fig. 6, how many current pulses will there be through the load (60-cycle power line)?
   (b) In the circuit shown in Fig. 9, how many current pulses will there be through the load (60-cycle power line)?

6. In the circuits shown in Fig. 11, what part supplies current to the load when the rectifier is not conducting?

7. Explain the following things in connection with the L-C circuit shown in Fig. 15:
   (a) The action of the choke when ac flows through it.
   (b) The action of the choke when dc flows through it.
   (c) The action of the capacitor when ac flows through it.
   (d) The action of the capacitor when dc flows through it.

8. What type of filter network has better voltage regulation -- the capacitor input or the choke input?

9. If you are servicing a five-tube table model radio that uses a universal ac-dc power supply and you see that two of the tubes are not lighting, where would you look for trouble?

10. How does an increase in output voltage affect Q1 and Q2 in the regulated power supply shown in Fig. 26?

HOW TO START STUDYING

For some people, starting to study is just as hard as getting up in the morning. An alarm clock will work in both cases, so try setting the alarm for a definite study-starting time each day. Start studying promptly and definitely, without sharpening pencils, trimming fingernails or wasting time in other ways.

Beginning is for many people the hardest part of any job they tackle. So formidable does each task appear before starting that they waste the day in dilly-dallying, in day-dreaming, and in wishing they didn't have to do it. The next day and the next after that are the same story. Indecision brings its own delays, making it harder and harder to buckle down to work.

Are you in earnest? Then seize this very minute; begin what you can do or dream you can. Boldness in starting a new lesson is a great moral aid to mastery of that lesson; only begin, and your mind grows alert, eager to keep on working. Begin, and surprisingly soon you will be finished.

nri

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-214

TELEVISION • RADIO SERVICING

POWER SUPPLIES FOR ELECTRONIC EQUIPMENT

B201

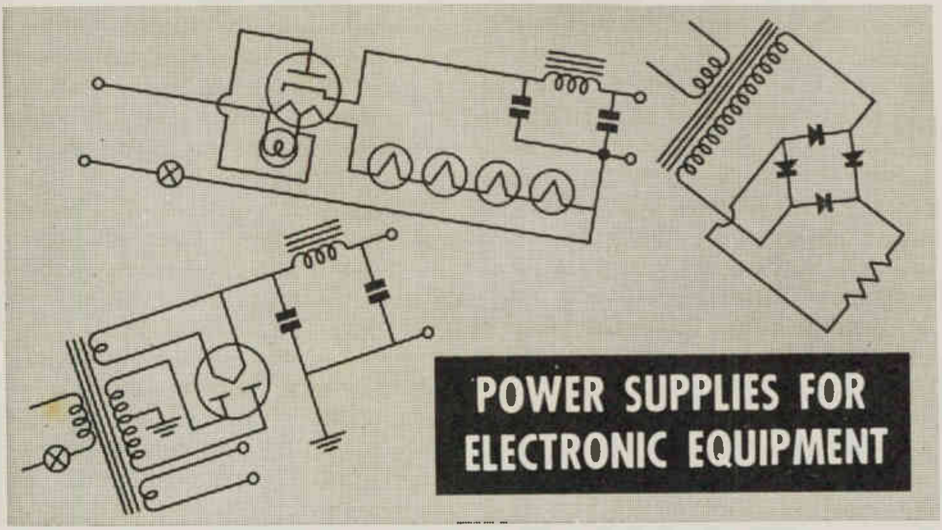
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

POWER SUPPLIES
FOR
ELECTRONIC EQUIPMENT

B201

STUDY SCHEDULE NO. 1

JDL5M369 **1969 EDITION** Litho in U.S.A.



POWER SUPPLIES FOR ELECTRONIC EQUIPMENT

One of the most important sections in electronic equipment is the power supply. It is the section that furnishes the operating voltages and currents required by the various stages. If the power supply is not operating properly, the equipment can't do the job it's supposed to do.

In your career as an electronics technician, you will encounter many different types of power supplies. In equipment using tubes you will find power supplies that must supply the heater voltage for the various tubes and, in addition, a dc supply voltage that is often considerably higher than the power line voltage.

On the other hand, the power supply in transistorized equipment will be quite different from that used in tube-operated equipment. Most power supplies in transistorized equipment will have to reduce the voltage to some value less than the line voltage.

However, the current requirements of the power supply in a transistorized piece of equipment may be considerably higher than those of a power supply in a similar tube-operated device. All power supplies have basically the same function, regardless of the parts and circuitry used to make them; the power supply must supply the operating voltages and currents required by the various stages in the equipment.

You have already studied the basic components used in power supplies. In this lesson you will learn more about these components and how they are used together in this particular application. You will be introduced to some new circuits and will learn enough about power supplies to enable you to understand the purpose for which each part in a power supply is used. Once you know why the various parts are used and understand what each one is supposed to do, you should be able to service any power supply defect you encounter.

We will first take up the different rectifier circuits used in modern power supplies. The power supplied by power companies for home and industrial use is ac power, whereas the tubes and transistors used in electronic equipment require dc operating voltages. Therefore, in a power supply designed to operate from a power line, we must have some means of changing the ac to dc. The device used to do this is called a rectifier.

Once the ac is changed to dc by a rectifier, we have what is called a pulsating dc at the output of the rectifier. This is actually dc with ac superimposed on it. A power supply must therefore have some means of filtering or smoothing the pulsating dc to get pure dc. This is done by means of a filter network, which separates the ac and dc components of the pulsating dc at the rectifier output so that only the dc appears at the output of the filter network.

Many power supplies also have some type of voltage-divider network. Such a network is designed to provide several different operating voltages from one power supply. All the tubes or transistors in a piece of electronic equipment may not require the same operating voltage. It is more economical to use a single power supply and a voltage divider than to use a separate power supply for each voltage needed.

The power supplies in modern electronic equipment use solid state rectifiers in most low-voltage applications. Vacuum tubes are seldom used today as rectifiers in such devices as radio or television receivers or in other modern equipment. However, there are still millions of radios and television receivers in use today that do use vacuum-tube rectifiers. Therefore, you will probably have to work on this type of power supply as a service technician even though it is obsolete as far as its use in new equipment is concerned and you still should know how this type of rectifier works. For this reason, we will cover not only the new rectifier circuits using solid state rectifiers, but also a number of the older frequently-used rectifier circuits using vacuum tubes.

Rectifier Circuits

Any device that will pass current in one direction but not in the other direction can be used as a rectifier. You have already seen one example of this type of device: the vacuum tube. In a vacuum tube, as long as the voltage applied to the plate is positive with respect to the voltage applied to the cathode, the current will flow from the cathode to the plate of the tube. However, if the voltage applied to the plate is negative with respect to the voltage applied to the cathode, there will be no current flow through the tube because current cannot normally flow through the tube from the plate to the cathode. Thus, a two-element or diode tube was used for many years as the rectifier in the power supply of radio and television receivers.

The diode tube is entirely satisfactory as a rectifier, but it does have one big disadvantage. In order to handle the currents required in large radio receivers or in television receivers a rectifier tube with a rather heavy cathode or filament is required. Considerable power must be applied to the heater to heat the large cathode or filament, thereby bringing it to the temperature required for it to emit an abundant supply of electrons. Not only does this increase the power consumed by the equipment; also, a substantial amount of heat is given off by the diode and this, in turn, heats up other parts in the equipment. This often contributes to a shortened life of the other parts.

As we mentioned, diode vacuum tubes were used for many years as the rectifiers in radio and television receivers. However, a number of years ago the selenium rectifier began to replace the vacuum tube as the rectifier in entertainment-type equipment.

A typical selenium rectifier is shown in Fig. 1. A selenium rectifier is made up of a series of selenium discs with a coating of selenium oxide on the surface of one side of each disc. Electrons can flow from the selenium to the selenium oxide quite readily, but they cannot readily flow in the other direction, from the selenium oxide to the selenium. Thus, a selenium rectifier permits current to flow through it in one direction, but offers a high resistance to current flow through it in the opposite direction.

This type of rectifier is often called a dry-disc rectifier: "dry" to distinguish it from earlier rectifiers that used a wet chemical solution, "disc" because it is made up of discs. The square plates that are visible in Fig. 1 are cooling fins. The discs used are usually round and are

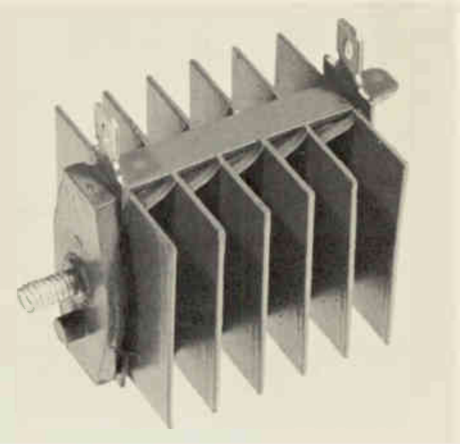


Fig. 1. A typical selenium rectifier designed for use in electronic equipment.

placed between the cooling fins, which are necessary because the rectifier does have some resistance and the current flowing through this resistance produces heat which must be dissipated.

The advantage of the selenium rectifier over the diode tube is that the selenium rectifier does not have a cathode that must be heated, and hence the power required to heat the cathode is saved. In addition, the total heat dissipated into the equipment from the selenium rectifier is somewhat lower than from a tube capable of handling the same current.

Both the vacuum tube rectifier and the selenium rectifier have been replaced in modern radio and television receivers by the silicon rectifier. The silicon rectifier, like the selenium rectifier, does not require any heater power; in addition, a silicon rectifier is much smaller than a selenium rectifier. It has a much lower forward resistance; that is, it offers far less opposition to current flow through it in the forward direction than does a selenium rectifier, and at the same time it has a higher reverse resistance (in other words it will permit a smaller current to flow through it in the reverse direction than a selenium rectifier).

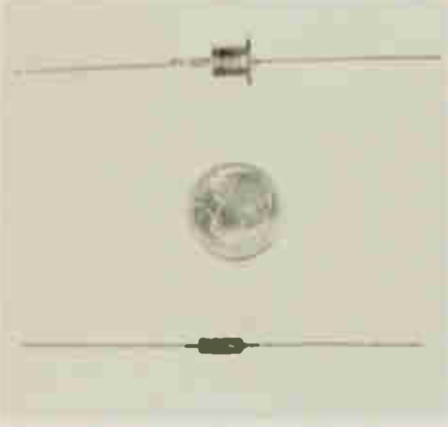


Fig. 2. Two typical silicon rectifiers with a dime between them to show their relative size.

Two typical silicon rectifiers are shown in Fig. 2. The rectifier at the top is called a top-hat rectifier because of its shape. A rectifier of this type and size is capable of handling currents several times those required in a color TV receiver. We have shown a photograph of the two rectifiers with a dime in between them so you can get an idea of the relative size of the two units. The lower rectifier is capable of handling currents of two or three amperes.

In addition to their small size and high current-handling capabilities, silicon rectifiers have another big advantage over selenium and vacuum tube rectifiers due to modern manufacturing techniques: they are relatively inexpensive to manufacture. Furthermore, unless they are overloaded, their life is almost indefinite.

Now, let's see how the various types of rectifiers are used in power supply circuits.

HALF-WAVE RECTIFIERS

You already know that the power supplied by most power companies is ac power and that the voltage supplied has a waveform that is called a sine wave. A typical single cycle is shown in Fig. 3A. A rectifier circuit using a diode tube is shown in Fig. 3B. Let's assume that the waveform shown at A represents the voltage at terminal A of Fig. 3 with respect to the voltage at terminal B. This means that for the first half-cycle (that is, the voltage waveform

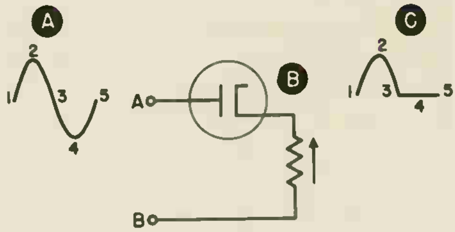


Fig. 3. How current flows in a half-wave rectifier circuit for one ac cycle. We have omitted the rectifier heater to simplify the diagram.

from point 1 to point 3), the plate of the tube will be positive. When the plate of the tube is positive, it will attract electrons from the cathode; therefore, current can flow through the tube.

Thus during the first half-cycle, as the plate voltage starts at point 1 and builds up to point 2, the current through the tube will increase from point 1 to point 2 as shown in Fig. 3C. As the voltage decreases during the first half-cycle from point 2 to point 3, the current through the tube will decrease as shown from point 2 to point 3 in Fig. 3C. When the voltage in Fig. 3A reaches point 3 there will be zero potential between points A and B in the rectifier circuit and current will stop flowing.

During the next half-cycle terminal A will be negative with respect to terminal B. This means that the plate of the tube will be negative; hence no current can flow through the tube. Therefore, the current will be zero (as shown in the waveform in Fig. 3C) as the voltage swings from point 3 to point 5.

The rectifier circuit shown in Fig. 3 conducts when the plate of the tube is positive. Since this occurs during one half the time of each cycle then the rectifier conducts during one half of the cycle but not during the other half. As a result, the rectifier is called a half-wave rectifier. If we operate this type of rectifier from a 60-cycle power line, we will get 60 current pulses through the rectifier during one half-cycle and 60 intervals during which there is no current flow through the rectifier.

Another half-wave rectifier is shown in Fig. 4. Here we have shown a solid-state rectifier in place of the tube. This could be either a selenium rectifier or a silicon rectifier -- the same symbol is used for both types.

Notice the schematic symbol used for the rectifier. Also notice that the arrows indicate that the direction of current flow through the circuit is opposite to the direction in which the arrow points in the schematic symbol. The reason for this is that in the early days of electricity, scientists thought that current flowed from positive to negative. Therefore, this symbol was designed to show the direction in which current flowed. But it was discovered later that current flow was actually electron flow and that it flowed from negative to positive, a direction opposite from that in which the early scientists thought it flowed. Although we know current flow is from negative to positive, we still use the same symbol; it has never been changed so that the arrow is actually pointing in the direction opposite to the direction of electron flow.

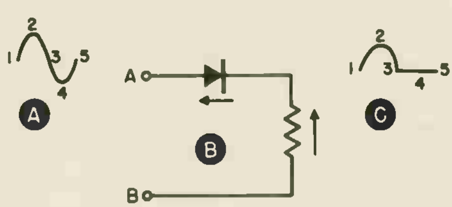


Fig. 4. A half-wave selenium rectifier circuit.

In a power supply of this type using a selenium rectifier, with terminal A positive with respect to terminal B, the selenium rectifier will offer only a low resistance to the flow of current through it; therefore, current flows in the circuit from B through the load and through the rectifier and back to terminal A. During the next half-cycle, when terminal A is negative with respect to terminal B, the selenium rectifier will offer a very high opposition to the flow of current through it so that there will be little or no current flow through the load (as shown in Fig. 4C).

We mentioned that the schematic symbol for the rectifier in Fig. 4 also represents a silicon rectifier. If a silicon rectifier is used, the rectifier will simply consist of a PN junction. The P-type material will be on the side represented by the arrow and the N-type material by the flat line. With a PN junction rectifier in the circuit when terminal A is positive and terminal B is negative, we will have a positive voltage applied to the P side of the junction and a negative voltage applied to the N side of the junction. The negative voltage will repel electrons from the N side of the junction across the junction into the P-type material. Electrons will be attracted through the P-type material by the positive potential applied to it. In other words, there will be a forward bias placed across the junction and current can readily flow through the rectifier because the carriers can cross the junction.

During the next half-cycle when the polarity reverses, terminal A will be negative and terminal B positive. Thus there will be a negative voltage, applied to the P side of the junction, which will repel electrons and prevent them from crossing the junction. At the same time there will be a positive voltage, applied on the N side of the junction, which will prevent any holes from the P side crossing the junction. In other words, there will be a reverse bias placed across the junction; hence the carriers cannot cross the junction and there will be no current flow through the circuit.

Of the three types of half-wave rectifiers, the silicon-type rectifier is the most widely used in modern equipment because of its small size, low cost and very low forward voltage drop.

Sometimes, in order to operate a piece of electronic equipment, a higher voltage is required than can be obtained directly from the power line. Under these circumstances a step-up transformer may be used to step up the voltage as shown in Fig. 5. The secondary-to-primary turns-ratio is simply adjusted to provide the required voltage step-up. A half-wave rectifier is then used as shown to rectify the ac and change it to pulsating dc. The operation of the half-wave rectifier is exactly the same in the circuit as

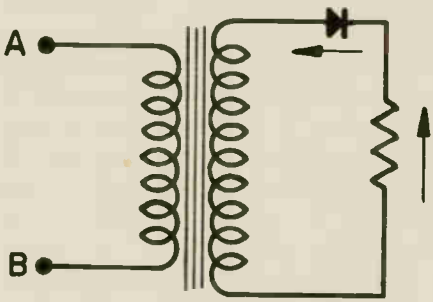


Fig. 5. A half-wave rectifier circuit using a power transformer to produce an output voltage greater than the power line voltage.

it was in the preceding circuits; however, the rectifier will have to have a higher voltage rating to make up for the fact that it is being used in a higher voltage circuit.

The disadvantage of the circuit shown in Fig. 5 is that power transformers are comparatively expensive. By means of the circuit shown in Fig. 6 a voltage approximately twice the voltage obtainable from the half-wave rectifier circuits shown in Figs. 3 and 4 can be obtained. This circuit can be operated directly from the power line and is known as a voltage-doubler circuit.

The operation of this circuit is quite simple. During one half-cycle terminal A will be negative with respect to terminal B. During this half-cycle electrons flow from A into the side of the capacitor C1 marked with the minus sign. The electrons flowing into this side of the capacitor force electrons out of the other side leaving a positive charge on this side of C1. The electrons leaving the positive side of C1 flow through the rectifier D1, back to side B of the power line which is positive and which will attract electrons. Thus, during this half-cycle, when terminal A is negative with respect to terminal B, capacitor C1 is charged with the polarity shown. The peak charge on C1 will be equal to the peak value of the ac input voltage.

During the next half-cycle, when terminal A is positive with respect to terminal B, we have a situation where the voltage between terminals A and B is effectively placed in series with the voltage charging capacitor C1. These series-connected voltages will cause a current to flow through the load and through D2. During this half-cycle terminal A is

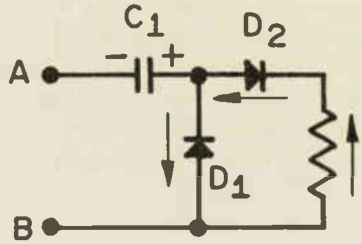


Fig. 6. A half-wave voltage-doubler circuit.

positive with respect to terminal B. This means that terminal B is negative. Electrons flow from terminal B through the load and through D2. They are attracted by a positive voltage which is equal to the voltage across C1 plus the line voltage. Thus the peak voltage that can be developed across the load will be equal to twice the peak line voltage.

You might wonder why the current flows through only one rectifier during each half-cycle. During the first half-cycle, when terminal A is negative with respect to terminal B, the electrons flowing through D1 and charging C1 cannot flow through D2 because the diode is connected in such a way as to prevent current flow through them in that direction. Similarly, during the next half-cycle, when terminal B is negative and terminal A is positive, current cannot flow through D1 because it would have to flow through it in the reverse direction. Current can flow through the diodes only in the direction shown and it will flow through D1 during one half-cycle and through D2 during the other half-cycle.

This type of power supply is known as a half-wave doubler circuit. It is called a voltage-doubler circuit because the voltage across the load is effectively double the line voltage.

It is called a half-wave circuit because there is a current pulse to the load during only one half of each cycle. The half-wave voltage-doubler circuit is widely used in modern radio and television receivers. It's a very important circuit and you should be sure you understand how it works before leaving it.

When a half-wave rectifier is used in the power supply, the current will flow through the rectifier in a series of pulses. With a 60-cycle power supply line, there will be 60 pulses per second: one pulse during each positive-half cycle and nothing during each negative-half cycle. The net result is that you will have current flowing through the rectifier for no more than half the time. This results in a pulsating dc output from the rectifier that is rather difficult to smooth out to the pure dc required in most equipment to operate the tubes and/or transistors. A somewhat better arrangement is the full-wave rectifier that passes current during both halves of the ac voltage cycles.

FULL-WAVE RECTIFIERS

A typical full-wave rectifier circuit is shown in Fig. 7. Notice that the tube used is a twin diode tube.

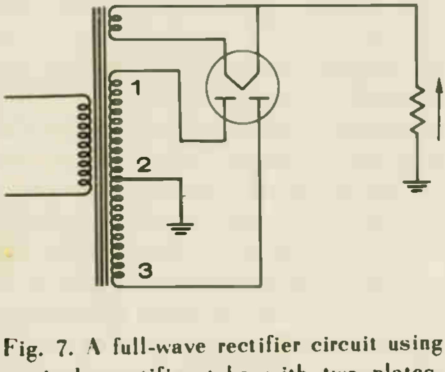


Fig. 7. A full-wave rectifier circuit using a single rectifier tube with two plates.

The tube has two plates and a single filament which is used with both plates. In operation, this tube acts as two separate diode tubes.

The power transformer used in the rectifier circuit has three windings. The primary winding is the winding that connects to the power line. A low-voltage winding is used to provide the current required to heat the filament of the rectifier tube. It serves no other purpose as far as the operation of the rectifier circuit is concerned. This winding is often referred to as the filament winding.

The high-voltage winding on the transformer is the winding that will supply the pulsating current to the load resistor. Notice that this winding has a center tap. In operation, one half of the winding first supplies the current and then, during the next half-cycle, the other half of the winding supplies current to the load.

We can see how this rectifier circuit works if we consider one half-cycle during which terminal 1 of the high-voltage secondary is positive. This means that terminal 2 will be negative with respect to terminal 1 and terminal 3 will be even more negative. Electrons will leave the center tap, terminal 2, and flow through ground to the load resistor. They will flow through the load resistor to the filament of the rectifier tube and then be attracted to the plate connected to terminal 1 because this plate has a positive voltage applied to it. No electrons will flow to the other plate because this plate is negative with respect to both terminals 1 and 2.

During the next half-cycle, the polarity of the secondary voltage will reverse. At this time terminal 3 will be positive, terminal 2 negative with respect to it, and terminal 1 even more negative. During this half-cycle electrons will leave terminal 2 and flow through ground to the load, through the load and to the filament of the rectifier tube, and then to the plate connected to terminal 3 because this plate now has the positive voltage applied to it. No electrons will flow to terminal 1 because terminal 1 is negative with respect to both terminals 2 and 3.

Notice that with the full-wave rectifier circuit we get a current pulse through the load resistor during each half-cycle. This means that for a 60-cycle power line we will get 120 pulses of current through the load. Since there is current flowing through the load during each half-cycle, this type of rectifier produces an output that is much easier to filter to a smooth dc than the output from a half-wave rectifier.

Either selenium rectifiers or silicon rectifiers can be substituted in this type of circuit in place of the rectifier tube. This type of circuit was widely used in older television receivers along with vacuum tubes. Modern TV receivers use silicon rectifiers and frequently use bridge-rectifier circuits or voltage-doubler circuits in place of this circuit.

BRIDGE-RECTIFIER CIRCUITS

One of the disadvantages of the full-wave rectifier circuit shown in Fig. 7 is that it requires a transformer with a center-tapped secondary. The total voltage across the entire secondary winding is actually twice the voltage between the center tap and either end of the secondary winding. This type of transformer is more expensive to manufacture than a transformer without a center tap because twice as many turns are required on the secondary to get the required voltage.

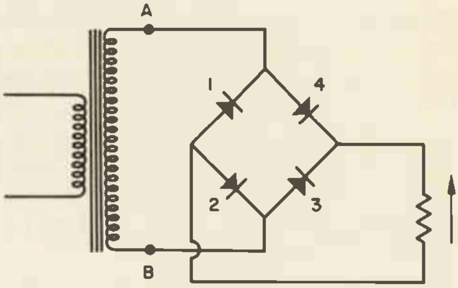


Fig. 8. A bridge-rectifier circuit.

A circuit that gets around the requirement of a center-tapped secondary is shown in Fig. 8. This is called a bridge-rectifier circuit; it is also often called a full-wave bridge-rectifier circuit because current flows to the load during each half-cycle.

A quick look at the circuit immediately shows us that four rectifiers are required in a circuit of this type. At one time this was a disadvantage because of the cost of rectifiers, but silicon rectifiers are comparatively inexpensive today and it is usually more economical to use the extra two silicon rectifiers and avoid the center tap on the secondary winding of the power transformer. The power transformer shown in Fig. 8 would be far more economical to manufacture than the one shown in Fig. 7.

The operation of the bridge rectifier is comparatively simple. When terminal A is positive and terminal B is negative, current will flow from terminal B through the rectifier marked 2 on the diagram and then through the load to the junction of rectifiers 3 and 4. It will then flow through rectifier 4 back to terminal A on the transformer. During the next half-cycle, when terminal A is negative and terminal B is positive, current will flow from terminal A on the transformer through the rectifier marked 1 on the diagram and then through the load back to the junction of rectifiers 3 and 4. This time the current will flow through rectifier 3 back to terminal B of the power transformer.

Notice that during each half-cycle current flows through two of the rectifiers. During one half-cycle it will flow through rectifiers 2 and 4 and during the other half-cycle it will flow through rectifiers 1 and 3. Also notice that current flows during both half-cycles; therefore, this bridge rectifier is a full-wave rectifier.

Bridge rectifiers have been used in television receivers where comparatively high operating voltages and high currents are required. The bridge circuit eliminates the need for the center tap on the power transformer secondary winding and thus reduces the cost of the transformer. The voltage regulation (the ratio of the full-load voltage to the no-load voltage) obtainable with this type of power supply is as good as the regulation that can be obtained from the full-wave rectifier circuit shown in Fig. 7.

FULL-WAVE VOLTAGE DOUBLERS

A full-wave voltage-doubler circuit is shown in Fig. 9. One advantage of this type of circuit is that we get the same load voltage as you have in a circuit like the bridge rectifier shown in Fig. 8, although only half as many turns are required on the secondary of the power transformer. This will result in a savings in the cost of the power transformer. Another advantage of this type of circuit is that only two rectifiers are required instead of the four required in the bridge-rectifier circuit.

The operation of the full-wave voltage-doubler circuit is quite simple. During one half-cycle terminal A of the power transformer secondary will be negative and terminal B will be positive. During this half-cycle current flows from terminal A through diode D2 to the capacitor C2 charging the capacitor as shown. Electrons flow into the negative side of the capacitor and out the positive side back to terminal B of the power transformer. During the next half-cycle, terminal B of the power transformer secondary will be negative and terminal A will be positive. During this half-cycle electrons leave terminal B of the power transformer and flow into capacitor C1. They flow into the side of the capacitor marked with the minus sign and force electrons out of the plus side. Electrons leaving the plus side flow through diode D1 back to terminal A of the power transformer, which is positive.

The capacitors C1 and C2 are connected in series and they supply the voltage to the load. The capacitors are charged by the current flowing

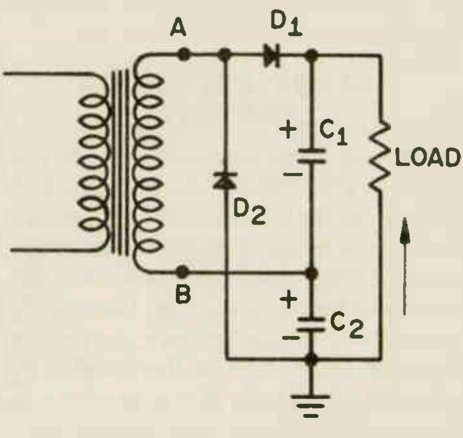


Fig. 9. A full-wave voltage doubler.

through the diodes, and since there is a charging pulse each half-cycle we will get 120 charging pulses from a 60-cycle power line. Since there is a charging pulse during each half-cycle, the circuit is a full-wave rectifier. The actual voltage that will be available across C1 and C2 in series will depend upon the power transformer secondary resistance, the resistance of the diodes when they are conducting, the size of the two capacitors, and the size of the load. As the resistance of the load increases, the current that will flow through the load decreases, and the charge on each capacitor becomes closer to the peak value of the ac voltage between terminals A and B.

Notice that in the diagram we have shown the direction of current flow through the load. The load current is supplied entirely by the charged capacitors C1 and C2. As the capacitors supply current to the load, electrons leave the negative plate of C2 and flow through the load in the direction shown. These electrons flow into the positive side of C1 forcing electrons out of the negative side into the positive side of C2. Thus the current flow through the load tends to reduce the charge across the capacitors. Of course, during each half-cycle one of the diodes conducts to build the charge across one of the capacitors up towards the peak value of the line voltage.

While this type of rectifier circuit offers some advantages over the circuits shown in Figs. 7 and 8, it does not have as good voltage regulation as they have. However, by using capacitors of large capacity for C1 and C2, and with modern silicon rectifiers that have a very low resistance when they are conducting, reasonably good voltage regulation can be obtained from this type of power supply. You will find the power supply widely used both in monochrome and color television receivers. Be sure you understand how it operates because you will run into it frequently.

In comparing this power supply with the half-wave doubler circuit shown in Fig. 6, we immediately see that the full-wave doubler circuit is best suited to equipment where a power transformer is used. The half-wave voltage-doubler circuit is widely used in equipment where no power transformer is used.

SUMMARY

The rectifier circuits that we have discussed in this section of the lesson are extremely important. As you know, the power supplied by the power companies is alternating current and the tubes and transistors in electronic equipment require direct current for their operation. Therefore, equipment designed to operate from the power line must use some type of rectifier to convert the alternating current to direct current. One of the circuits shown in this section of the lesson is likely to be found in any type of electronic equipment you will service.

The half-wave rectifier circuit shown in Fig. 5 is perhaps the most widely used. All table model radio receivers use this type of rectifier circuit without a power transformer so that the receiver can operate directly from the power line. The half-wave voltage-doubler circuit shown in Fig. 6 is widely used in television receivers where operating voltages higher than those that can be obtained directly from the power line are needed. In some of the older radio receivers and many older television receivers a full-wave rectifier circuit such as shown in Fig. 7 will be found. This type of circuit was used almost exclusively in television receivers before the development of low-cost selenium and silicon rectifiers.

The bridge-rectifier circuit shown in Fig. 8 is used in many television receivers, particularly large television receivers where fairly high voltages and high currents are required. This type of circuit is also used in some transistorized equipment where lower than line voltages are required. In this instance, instead of being a step-up transformer, the transformer will be a step-down transformer that steps the line voltage down to a low value. The bridge-rectifier circuit is then used so that good voltage regulation can be obtained and so that at the same time a comparatively large current can be taken from the power supply.

The full-wave voltage-doubler circuit has been used in many television receivers -- in those which use power transformers where voltages higher than the line voltage are required -- yet at the same time, the transformer serves primarily as an isolation transformer. In other words, the secondary voltage is approximately equal to the primary voltage. The higher voltage required is obtained by the voltage-doubling action. The output from the full-wave doubler circuit is somewhat easier to filter or smooth out than the output from the half-wave doubler circuit, because with the former we have 120 pulses per second through the load and with the latter we have only 60 pulses per second.

Before leaving this section regarding rectifier circuits, why not try to draw the circuits yourself? It would be worthwhile, and you don't have to draw them from memory -- copying them first from the book will help you to remember what they look like. Eventually, you'll be able to draw them from memory and recognize various circuits on the schematic diagram of any radio or TV receiver or piece of electronic equipment you may encounter.

After reviewing this section, do the following self-test questions.

SELF-TEST QUESTIONS

(a) In a half-wave rectifier circuit such as the circuit in Fig. 5, how many current pulses per second will there be through the load when the power line frequency is 60 pulses per second?

(b) What is the disadvantage of a half-wave rectifier circuit?

(c) What is the purpose of the diode marked D1 in the half-wave voltage-doubler circuit shown in Fig. 6?

(d) Why is the circuit shown in Fig. 7 called a full-wave rectifier circuit?

(e) What is the disadvantage of the full-wave rectifier circuit shown in Fig. 7?

(f) What are the advantages of the bridge-rectifier circuit shown in Fig. 8?

(g) What advantage does the voltage-doubler circuit shown in Fig. 9 have over the voltage-doubler circuit shown in Fig. 6?

(h) What are the advantages and disadvantages of the full-wave voltage doubler over the bridge-rectifier circuit?

Filter Circuits

The output from the rectifiers we discussed in the preceding section is not pure dc. Instead, it is pulsating dc: "direct" because it flows in only one direction, "pulsating" because it is varying in amplitude rather than flowing steadily. A pulsating dc voltage is a voltage that does not change polarity, but does change in amplitude. The voltage at the output of a half-wave rectifier will be zero during one half of each cycle and swing in a positive direction during the other half of each cycle.

Looking at the half-wave circuits shown in Fig. 3 and 4, you can consider the rectifier more or less as a switch. During one half-cycle the switch is closed so that the load is connected directly across the power line, and during the other half-cycle the switch is open so that no voltage is applied to the load. The voltage across the load in a half-wave rectifier circuit looks like Fig. 10A. The first half-cycle represents the cycle when the switch is closed and the load is connected directly across the power line, and the second half-cycle represents the cycle when the switch is open and there is no voltage applied across the load. We have shown what the voltage across the load will look like for four cycles in Fig. 10A.

In a full-wave rectifier circuit such as shown in Fig. 7, you have two switches. During one half-cycle one switch closes and connects the load across one half of the power transformer secondary; then, during the next half-cycle, the other switch closes and connects the load across the other half of the transformer secondary.

In the bridge rectifier shown in Fig. 8 two rectifiers act as switches and close to connect the load across the transformer secondary during one half-cycle; then during the next half-cycle the other two switches close, giving the effect of turning the load around so that the current flows through it in the same direction and the voltage applied across the load has the same polarity. The output from a full-wave rectifier circuit will produce a voltage across the load that looks like Fig. 10B.

As you can see from the waveforms shown in Fig. 10, the output taken directly from the rectifier is not a pure dc. There is a voltage, but the voltage drops to zero, builds up to the maximum value, and drops to zero again. In the half-wave circuit it remains at zero for a half-cycle and then builds up again in a positive direction. In the full-wave rectifier circuit the voltage builds up across the load during each half-cycle. In either case, this pulsating voltage will cause a pulsating current through the load which is entirely unsuitable for use in electronic equipment. Fortunately, there are convenient methods that can be used to filter or smooth this voltage to a pure dc voltage.

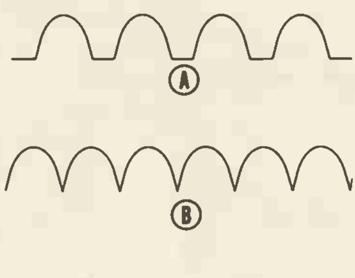


Fig. 10. Output voltage from half-wave and full-wave rectifier.

The pulsating dc voltage at the output of the rectifier is actually a dc voltage with an ac voltage, called a ripple voltage or a hum voltage, superimposed on it. The circuits used to get rid of this ripple or hum voltage are called filter circuits. There are a number of different types of filter circuits found in electronic equipment; in this section we will cover some of the circuits more commonly used.

THE SIMPLE CAPACITOR CIRCUIT

One of the simplest filters is the single capacitor filter shown in Fig. 11. In Fig. 11A we have shown a rectifier circuit using a tube and in Fig. 11B a rectifier circuit using a silicon rectifier. Notice that the circuits are practically identical -- we simply changed the rectifying devices in the two circuits.

The simple capacitor-type filter is sometimes used in circuits where the current drained or taken from the power supply is low. If the rectifier must supply high current to the circuit, this type of filter is generally unsatisfactory because there will be too much ripple or hum present across the load. In other words, the simple filter is simply not capable of eliminating all the ac or ripple voltage present at the output of the rectifier.

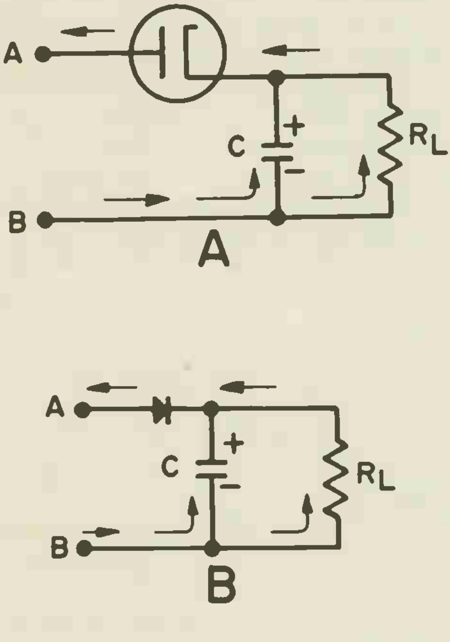


Fig. 11. A simple capacitor-type filter.

Both circuits shown in Fig. 11 work in the same way. Current flows through the rectifier during one half-cycle, as in the half-wave rectifier circuits we studied previously. When terminal A is positive, electrons will flow from terminal B through the load and through the rectifier back to terminal A. At the same time, electrons will flow into the negative side of the capacitor and out the positive side and through the tube or silicon rectifier back to terminal A. The capacitor eventually will be charged to a value almost equal to the peak line voltage. This will happen when the ac line voltage reaches its peak value with terminal A at its peak positive voltage with respect to terminal B.

Now, if the load on the rectifier circuit is light (that is, if the load resistor is a high resistance that draws very little current), as the ac input voltage between terminal A and B drops, capacitor C will begin to supply the current required by the load. Electrons will start to leave the negative side of the capacitor and flow through the load resistor back to the positive side of the capacitor. They will continue doing this as the ac voltage drops to zero and remains at zero during the next half-cycle and starts to build up again in the positive direction. The capacitor will continue to supply current to the load as long as the voltage across the capacitor is greater than the ac input voltage. Eventually, the input voltage will reach a value greater than the capacitor voltage; then we'll get a current flow into the capacitor and through the rectifier to recharge the capacitor.

In Fig. 12 we have shown the ac input voltage. Notice that during the first half-cycle between points 1 and 2 in Fig. 12A the ac voltage is increasing in a positive direction. Let's assume that terminal A in Fig. 11 is becoming positive with respect to terminal B. This explanation applies to both of the two circuits shown. During this first half-cycle the capacitor is charging and follows a curve as shown from point 1 to point 2 in Fig. 12B. Now, as the ac cycle drops from point 2 to point 3 on curve A in Fig. 12, the voltage drops faster than the capacitor discharges. During the interval from point 2 to point 5 and almost to point 6, as shown in Fig. 12A, the capacitor discharges very little. The discharge is shown on the curve from point 2 over to the number 5 on curve B. At this point the ac input voltage exceeds the capacitor voltage, so the capacitor is recharged again.

In circuits where the drain or load current is low, the capacitor will discharge very little between current pulses that recharge it so that the voltage across the capacitor and hence the voltage across the load remain almost constant. Of course, as the requirements of the load increase, the capacitor will discharge more so that there will be more of a voltage drop across the capacitor and across the load than in circuits where current drain is low.

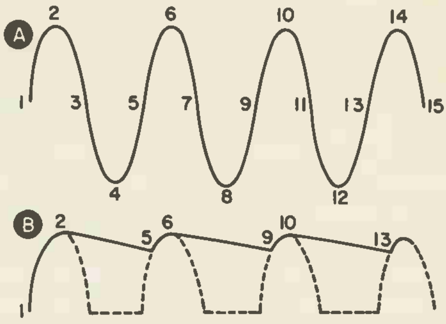


Fig. 12. Voltage waveshapes for a simple capacitor filter. (A) Input voltage; (B) output voltage.

There are several important points you should notice in the circuits shown in Fig. 11. Notice that the current does not flow during an entire half-cycle, but flows only when the line voltage exceeds the voltage across the capacitor. This may be for a very short interval if a load is a high resistance and draws very little current, or it may be for a sizable portion of a half-cycle if the current is a low resistance and draws a high current from the power supply. However, since the current flows through the rectifier in pulses, then the current pulse through the rectifier must be many times the average dc current flowing through the load. This is because the pulse or current that flows through the rectifier during the interval in which the rectifier is conducting must supply enough current to the capacitor to charge the capacitor and make up for the current it is going to supply for the remainder of the cycle.

When the rectifier is not conducting it is because the voltage across it is what we call a reverse voltage. In other words, it has a polarity opposite from that which the rectifier needs to conduct. In the case of the vacuum tube circuit this means that the plate of the tube is negative with respect to the cathode; in the case of the silicon rectifier, that there is a reverse bias across the junction.

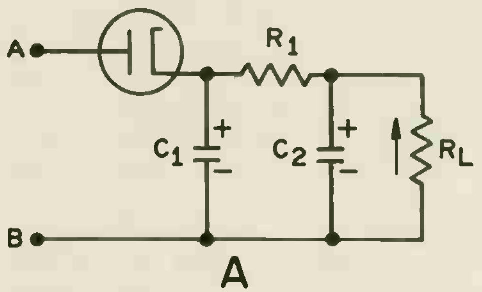
One of the important characteristics of a rectifier is the maximum peak reverse voltage that can be placed across the rectifier before it breaks down. In the circuit shown in Fig. 11 the capacitor will be charged as shown and the charge can equal the peak line voltage. During the next half-cycle, when the polarity of the input voltage reverses, terminal A will be negative and terminal B will be positive. When this voltage reaches its peak, the peak reverse voltage across the rectifier will be equal to twice the peak line voltage. The rectifier must be able to withstand this voltage without breaking down. This important characteristic, by which rectifiers are rated, is usually referred to as "PRV" (peak reverse voltage), although it may also be called "PIV" (peak inverse voltage). The two are simply the maximum reverse or inverse voltages that can be applied across the rectifier without its breaking down. In circuits such as those shown in Fig. 11, the PIV should be considerably higher than twice the peak line voltage in order to allow a reasonable safety factor.

As we mentioned previously, the simple capacitor-type filter shown in Fig. 11 is usable only where a small current is required by the load. If the current required is small, the output capacitor can be made large enough so that it discharges very little between pulses. If the current required by the load is high, on the other hand, then the capacitor will discharge appreciably between charging pulses, resulting in a varying voltage applied to the load. This is essentially the same as applying dc mixed with ac to the load. Additional filtering is required in applications of this type in order to eliminate ac so that we will have pure dc across the load.

AN R-C FILTER

An improved filter, which is often called a pi filter because it looks like the Greek letter pi (π), is shown in Fig. 13. You will notice that this filter consists of two capacitors, C1 and C2, and a filter resistor, R1.

The operation of the half-wave rectifier and capacitor C1, which is called the input filter capacitor, is the same as in the simple capacitor filter shown in Fig. 11. The rectifier tube passes current pulses to charge capacitor C1 with the polarity indicated on the diagram. However, if the



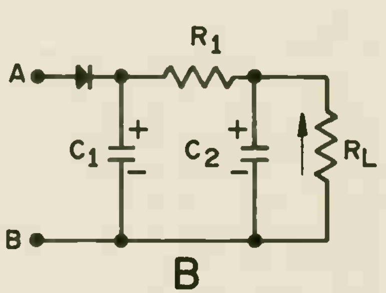


Fig. 13. An R-C pi-type filter.

load resistance R_L is low enough to draw appreciable current from the supply, then the voltage across C1 will discharge appreciably during the portion of the cycle when the rectifier tube is not conducting. Thus, we have dc with an ac superimposed on it across C1.

Now, to see the action of R1 and C2, let us first consider how the capacitor C2 reacts to ac and to dc. Remember that a capacitor is a device that will not pass dc -- it can be charged so that a dc voltage will exist across it, but applying dc to the plates of the capacitor will not cause a current to flow through it. Although electrons cannot cross through the dielectric of a capacitor, however, applying ac to the dielectric of a capacitor yields the effect of a current flowing through it. This is due to the fact that electrons will flow first into one plate and then into the other as the polarity of the ac voltage reverses.

You will remember from your study of capacitors that a capacitor offers what is called capacitive reactance (or opposition) to the flow of ac through it. The exact reactance that any capacitor will offer to the flow of ac through it is given by the formula:

$$Xc = \frac{1}{6.28 \times f \times C}$$

You can see from this formula that the larger the value of the capacitor, the lower the capacitive reactance will be to an ac voltage of a particular frequency. In Fig. 14A we have shown how the filter consisting of R1 and C2 reacts to dc; in Fig. 14B, we have shown how it reacts to ac.

As you know, a "perfect" capacitor will not pass dc -- although no

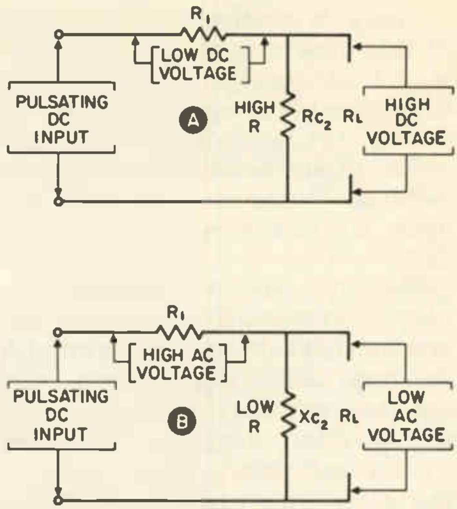


Fig. 14. Equivalent circuits showing the reaction of an R-C filter to dc at A, and to ac at B.

capacitor is really perfect because there will always be a small dc current (called a leakage current). In other words, the capacitor offers a very high resistance to dc, so we have shown it as resistor R_{c2}. Most of the dc voltage applied to the input of this filter network will appear across the capacitor, and there will be very little dc dropped across the resistor R1. The exact drop across this resistor will depend upon the size of the resistor and the current drawn by the load.

Now look at Fig. 14B which shows the reaction of the circuit to an ac voltage. The capacitor has a very low reactance to ac, so most of the ac voltage will be dropped across R1, because the resistance of R1 is much higher than the reactance (X) of C2. R1 and C2 act as a voltage divider network with most of the ac being dropped across R1 because its resistance is much higher than the reactance of C2.

There is another way of looking at this type of power supply which may help you see exactly what is happening in the circuit. Refer back to Fig. 13; the explanation applies to the circuit using the vacuum tube rectifier shown at A and to the one using the silicon rectifier shown at B.

When terminal A is positive with respect to terminal B, the diode will conduct and current can flow through the diode and through the load. During this part of the cycle electrons will flow into the plates of C1 and C2 marked with a minus sign. At the same time electrons will flow out of the other plate of both capacitors. Electrons leaving the positive side of C1 will flow directly through the diode being attracted by the positive voltage at terminal A. Because the resistance of the rectifier is low, C1 can charge up to a value almost equal to the peak ac voltage. However, the electrons leaving the positive side of C2 must flow through the filter resistor R1. Thus, capacitor C2 cannot charge to as high a voltage as capacitor C1.

When the ac input voltage drops so that the diode no longer conducts, capacitors C2 and C1 begin supplying the power required by the load. However, capacitor C1 is charged to a higher voltage than capacitor C2. Hence, capacitor C1 begins supplying power to the load and also tries to charge capacitor C2. Since electrons flowing from the negative side of C1 to the positive side must flow through filter resistor R1, the attempt of these electrons to charge C2 and flow through the load resistor will be somewhat restricted by R1. The net effect is that the resistor R1 prevents C2 from charging to as high a peak voltage as it would if R1 were not on the circuit. Because C1 charges to a higher voltage than C2, and while discharging tends to charge C2, the voltage across C2 is more nearly constant than the voltage across C1.

AN L-C FILTER

The disadvantage of the resistor-capacitor type of filter shown in Fig. 13 soon becomes apparent if you consider the size of the resistor and capacitor needed to obtain effective filtering and also the effect that the high value of resistance in the circuit has on the dc voltage present.

Consider Fig. 13 again for a minute. Suppose that the ac component of the pulsating dc across C1 is 10 volts and that the maximum ac component that can be applied to the load is only 1 volt. This means that the filter network consisting of R1 and C2 must produce a 9-to-1 voltage division. In other words, of the 10 volts ac appearing across C1 we must drop 9 volts across R1 and 1 volt across C2. This means that the resistance of R1 must be about 9 or 10 times the reactance of C2.

A 25-mfd capacitor (which is fairly large, particularly if it must be built to withstand high voltages), has a reactance of about 100 ohms at a frequency of 60 cycles. If we used such a capacitor for C2, then the resistance of R1 would have to be 10 times its reactance, or about 1000 ohms. If the current drawn by the load is 100 milliamperes, then the voltage drop across R1 due to the load current flowing through it will be 1000 ohms times .1 amp (100 ma), or 100 volts. This means that we will be losing 100 volts of our dc voltage across R1. Furthermore, the power being wasted by this resistor will be

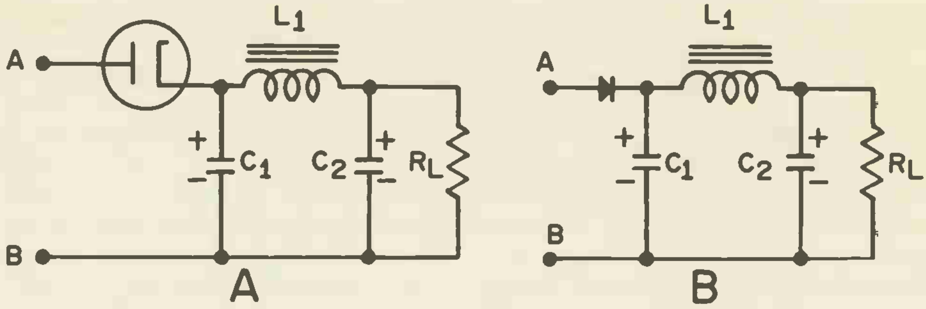


Fig. 15. An L-C filter.

equal to the voltage across it times the current flowing through it, which is 100 × .1, or 10 watts. You can see that we have an appreciable voltage drop and a sizable amount of power wasted by this resistor.

In some cases the current drawn through the filter resistor is not so high that the voltage drop across the resistor cannot be tolerated. Thus you will see this type of filter used in equipment when the current taken from the power supply is moderate. In equipment when the current drawn is high, a different type of filter is used. Such a filter is shown in Fig. 15. Notice that this circuit is identical to Fig. 13, except that we have substituted an iron-core choke for R1. The action of this filter network is quite similar to that of the filter network shown in Fig. 13. Again, capacitor C1 is charged by the rectifier and because there is no resistance in the circuit other than the rectifier resistance, it charges to a fairly high value. Just as before, however, the voltage across it will be pulsating (the equivalent of dc with ac superimposed on it).

Now consider the reaction of the choke to ac and dc. You will remember that a choke has inductance, and an inductance offers reactance to the flow of ac through it. At the same time the only opposition that the choke will offer to the flow of dc through it is due to the resistance of the wire used to wind the coil. By using a large size wire, this resistance can be kept quite low. The choke may have a dc resistance of 100 ohms or less and at the same time have a reactance of several thousand ohms to the 60-cycle ac applied to it.

The reaction of the L-C filter to dc is shown in Fig. 16A, and the reaction to ac is shown in Fig. 16B. In 16A we see that as far as the dc is concerned, the choke acts as a

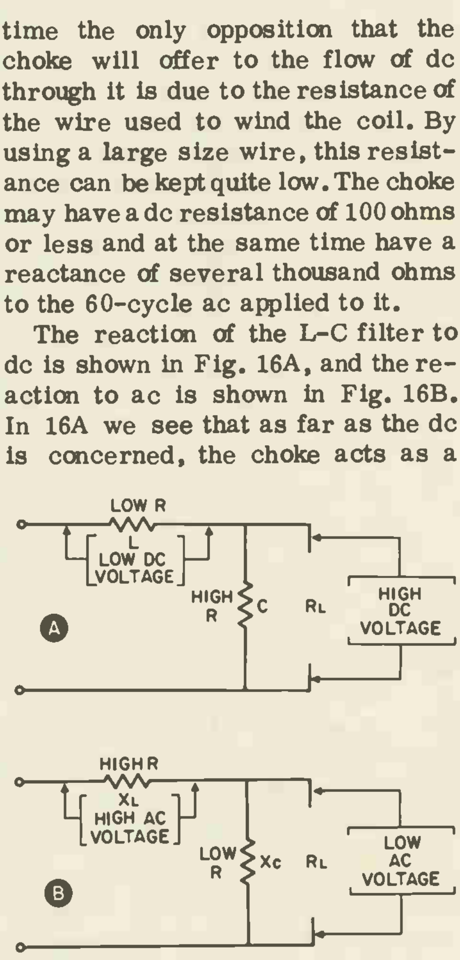


Fig. 16. Equivalent circuit showing the reaction of an L-C filter to dc at A, and to ac at B.

low resistance while the capacitor acts as a very high resistance. Thus, we have practically all of our dc voltage appearing across the capacitor and very little of it being lost across the choke. However, as shown in 16B, the choke acts as a high resistance to the ac while the capacitor acts as a low resistance. Thus, most of the ac voltage will appear across the choke and very little of it will appear across the capacitor.

There is another way of looking at the action of the L-C filter. We can consider the charging of the capacitors and the opposition offered by the choke more or less in the same way as we considered the action of the R-C filter in Fig. 13. During the first half-cycle, when terminal A of either rectifier circuit is positive and terminal B is negative, the rectifier will conduct. Electrons will flow from terminal B into the sides of C1 and C2 marked with a minus sign. These electrons will force electrons out of the positive side of C1 and they will flow through the rectifier with little or no opposition. However, the electrons leaving the positive side of C2 will encounter opposition in the choke. This is because at the instant when the electrons first try to get through the choke there is no magnetic field built up in the choke. You will remember that a choke is a device that opposes any change in current flowing through it. Therefore, the choke tries to keep the electrons leaving the positive side of C2 from flowing through it. Eventually this opposition offered by the choke is overcome, a magnetic field is built up in the choke, and some of the electrons can flow through the choke and capacitor C2 will be charged. However, the voltage to which C2 is charged will not be as high as the voltage to which C1 is charged because of the opposition of the choke.

When the ac input voltage drops below the voltage to which C1 is charged, the rectifier will no longer conduct and no current can flow through it. Now C1 and C2 and choke L1 must supply the current needed by the load. C2 does this by attempting to discharge. C1 also tries to discharge to charge C2 and to supply part of the current required by the load. At the same time there is a magnetic field built up in the choke L1 which does not collapse instantly but instead tries to keep current flowing in the direction it was flowing when the rectifier was conducting. It too helps to maintain the current flow through the load RL. Thus in this type of filter we have energy stored in three places: the capacitors C1 and C2 (as we did in the circuit in Fig. 13), and in the magnetic field of choke L1.

In the circuits shown in Fig. 15, the rectifier tube in Fig. 15A offers a certain amount of resistance to the flow of current through it. In addition, the rectifier tube has a cathode which must be heated by a heater. It takes some time for the cathode to come up to operating temperature; when a tube first starts conducting, the cathode is below normal operating temperature, and the tube offers a higher resistance to the flow of current through it than it does when the tube reaches its full operating temperature. Thus when the power supply is first turned on and the tube reaches a temperature at which the cathode begins to emit electrons, the tube offers considerable resistance to the flow of current through it. This limits the charging current through the tube that charges the input filter capacitor C1. In a matter of a few seconds capacitor C1 is charged and from then on the current that must flow through the tube is within the tube's capabilities.

In the circuit shown in Fig. 15B, however, the silicon rectifier does not have a cathode which must be heated -- as soon as the power is turned on the rectifier begins to conduct to charge the input capacitor C1. If you turn the equipment on at the peak of the ac cycle there will be a very high voltage immediately impressed across C1 and a very high current will flow through the rectifier. As a matter of fact, the current might be so high that it could burn out the rectifier. Even if the power is turned on when the ac voltage is at zero, a high current will flow through the diode to charge C1 as the voltage rises to a peak value with terminal A positive with respect to terminal B. If C1 is large enough, this could burn out the rectifier.

This problem of excessively high charging current can be overcome with the circuit shown in Fig. 17. Here a resistor is connected in series with the silicon rectifier to limit the current flow. In some equipment, this resistor is a fairly low resistance, fixed-value resistor. In other applications, the resistor may be a thermistor. You will remember that a thermistor is a resistor with a negative temperature coefficient. This means that the resistance of the thermistor decreases as its temperature increases.

With a thermistor for the resistor R1 shown in the circuit in Fig. 17, the thermistor will offer a fairly high resistance to current flow when the equipment is first turned on. This will limit the charging current that flows through the diode to charge C1 to a reasonable value. As the current flowing to the thermistor heats the thermistor and its resistance goes down, the charge across capacitor C1 will build up slowly so that the diode current never reaches an excessively high value. By the time the capacitor is fully charged, the thermistor temperature will have increased to a point where the resistance of the thermistor has dropped to a low value so that the thermistor has very little effect on the overall operation of the circuit.

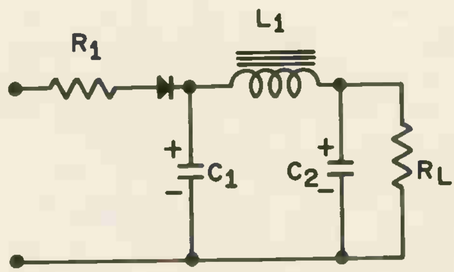


Fig. 17. Power supply with series-limiting resistor.

If you have to service a power supply of this type (where a resistor or a thermistor is used in the circuit in series with a silicon rectifier) and you find that the resistor or thermistor is opened, do not simply short the resistor or thermistor out of the circuit. If you do the chances are that the diode will burn out either the first time that you turn the equipment on or shortly thereafter.

The rectifier tubes used in low power equipment such as radio and television receivers are high vacuum rectifier tubes. However, in transmitters and in industrial applications where high voltages are involved you will often run into mercury-vapor rectifier tubes. Tubes of this type cannot be subjected to high peak currents through them without damaging the tube. A mercury-vapor rectifier tube in the circuit such as shown in Fig. 15A lasts only a short time. We can keep the peak current through the tube down to a safe value by using a somewhat different filter circuit known as a choke-input filter.

CHOKE-INPUT FILTERS

The filter circuits shown in Figs. 15 and 17 are called capacitor-input filters because the rectifier is connected directly to the input filter capacitor. A choke-input filter such as those frequently used with a mercury-vapor rectifier tube is shown in Fig. 18. Here, the rectifier tube is connected to a filter choke rather than a capacitor. Power supplies of this type will be found in radio and television transmitters and in industrial applications where comparatively high voltages are encountered.

It is easy to see how this type of filter works when we remember that the pulsating dc at the output of the rectifier tube is actually dc with ac superimposed on it. The choke offers little or no opposition to the flow of dc through it. On the other hand, the choke offers a high reactance to the flow of ac through it and at the same time the capacitor offers a low reactance to the ac across it. Thus the choke and the capacitor form a voltage divider network for the ac, so that most of the ac is dropped across the filter choke and very little of it appears across the load.

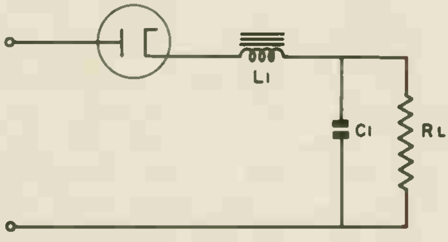


Fig. 18. A choke-input filter.

A more elaborate filter network is the two-section filter shown in Fig. 19. Again, it is a choke-input filter because the first element in the filter network is a choke. This type of filter network is frequently used in power supplies of radio and TV transmitters and of industrial electronic equipment where mercury-vapor rectifier tubes are used.

The choke-input filter has several advantages over the capacitor-input filter, even though the voltage obtained at the output of a choke-input filter is not quite as high as it is at the output of equivalent capacitor-input filters. In other words, if you feed the same pulsating dc into a choke-input filter you will not obtain as high an output voltage for a given load as you can with a capacitor-input filter. However, this type of filter has better voltage regulation than a capacitor-input filter. The voltage regulation of a power supply is the ratio of the full-load voltage to the no-load voltage. With a choke-input filter there is not as great a variation between the no-load and the full-load voltages as there is in the capacitor-input filter.

Another big advantage of the choke-input filter is that the peak current passed by the rectifier tube is held to a reasonable value. In a choke-input filter, the choke offers considerable reactance to any change in current flow through it. Thus, when the rectifier tube tries to conduct current heavily to charge C1 in Fig. 19, the input-filter choke L1 offers considerable reactance or opposition to the change in current flow through it. It tends to smooth

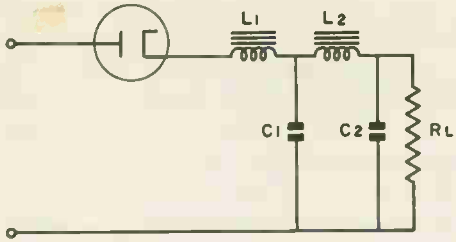


Fig. 19. A two-section choke-input filter.

out the pulses of current through the rectifier tube. The current pulse flowing through the rectifier tube flows for a slightly longer time than it would flow with an equivalent capacitor-input type of filter, and the peak amplitude of the current flowing through the rectifier tube is not as high as it would be with the capacitor input filter. This is a big advantage in a power supply using mercury-vapor rectifier tubes, because they can easily be destroyed by excessively high current pulses through them.

The filter network shown in Fig. 19 is quite effective in eliminating hum. Consider what would happen if the output of the rectifier had an ac voltage of 100 volts superimposed on the dc. If the two sections of the filter network are designed so that each choke has a reactance about 10 times as high as the reactance of the capacitors, each section will have approximately a 10-to-1 ripple voltage division; thus if there are 100 volts ac at the output of the rectifier, L1 and C1 will divide this voltage so that there will be only 10 volts appearing across C1. Now L2 and C2 act as a voltage divider network and divide this 10 volts further, so that the voltage across C2 would be only 1 volt. Thus, the two-section filter has reduced the ac hum or ripple voltage at the rectifier output from 100 volts to 1 volt.

The input filter choke, which is L1 in Fig. 19, is often a swinging choke. A swinging choke is designed so that it saturates rather easily and thus its inductance will vary appreciably as the current through the choke changes. When the current through the choke becomes high, the inductance and hence the inductive reactance of the choke decrease, but when the current through the choke is low, the inductance and hence the inductive reactance increase. Thus we have in effect a variable reactance between the rectifier tube and the input-filter capacitor; this variable reactance helps to improve voltage regulation at the power supply output. This type of choke is particularly useful in circuits where the load current goes through wide variations. If the load current goes down, the reactance of the choke increases. The increased reactance limits the charging action of the rectifier tube and keeps the output voltage from rising appreciably. On the other hand, if the load current increases, the reactance of the choke decreases, allowing the rectifier to charge C1 to a higher value so the capacitor can supply the increased current demand.

FACTORS AFFECTING THE OUTPUT VOLTAGE

In any filter network containing a filter choke or a filter resistor, the dc current flowing through the load must also flow through the filter choke or filter resistor. Thus, there will be a voltage drop across this choke or resistor; the exact value of the voltage drop will depend upon the dc resistance and the current flowing. We already pointed out that if a 1000-ohm filter resistor is used in the circuit and the current that is flowing is 100 milliamperes, the voltage drop across the filter resistor will be 100 volts. On the other hand, if a filter choke having a dc resistance of 100 ohms is used, a current of 100 milliamperes will produce a voltage drop of only 10 volts across it.

If the current drawn by the load changes, the output voltage at the output of the filter network will change. You can see why this is so--the current must flow through the filter resistor or filter choke. If the current flowing through the choke or resistor changes, the voltage drop across it will change, and hence the voltage at the output of the power supply will have some tendency to change.

The output voltage is also affected by the size of the filter capacitors used. If the filter capacitors used are small, then the rectifier is unable to keep them completely charged; if the filter capacitors are large, once the rectifier gets them charged, they will stay charged to a voltage near the peak ac voltage being applied to the rectifier tube.

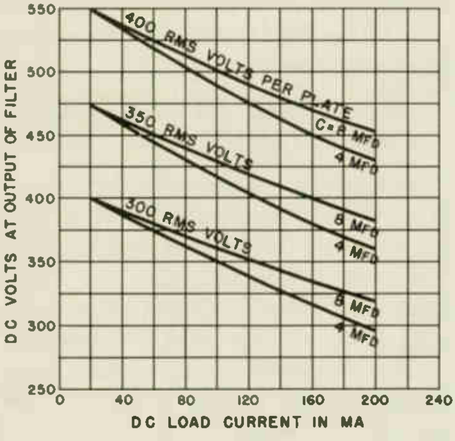


Fig. 20. How the size of the input capacitor of a filter affects the output voltage for different values of load current.

Of course, it is difficult to say whether a filter capacitor is large or small. Whether or not it is large for the particular circuit depends upon how much current is being drawn from the circuit. If the current drain is low, then a capacitor of 10 or 20-mfd will usually be sufficient to keep the power supply output voltage at or near the peak value of the ac voltage applied to the rectifier tube. However, if the current drain from the power supply is large, then the voltage across the power supply output will be considerably less than the peak ac value if the capacitors used are 10-mfd or 20-mfd capacitors.

To give you some idea of how the size of the input filter capacitor affects the output voltage, we have shown a graph in Fig. 20. The graph is for a full-wave rectifier. It shows how the output voltage varies with different load currents for both 4-mfd and 8-mfd capacitors. We have shown this for three separate input voltages. Notice that in each case when the current drain is low, the output voltage is substantially above the rms input voltage applied to the rectifier. Remember, of course, that the peak value of a 300-volt rms voltage is actually 1.4 times 300 volts.

Fig. 21 shows a comparison between a choke-input and capacitor-input type of filter. Notice that with a capacitor input the output voltage at low loads is substantially higher than it is for a choke input. However, as the load is increased, the output voltage from a capacitor-input type of filter drops rapidly. The output voltage from the choke-input type filter also drops as the load is increased, but it does not drop nearly as rapidly as it does with a capacitor-input type of filter.

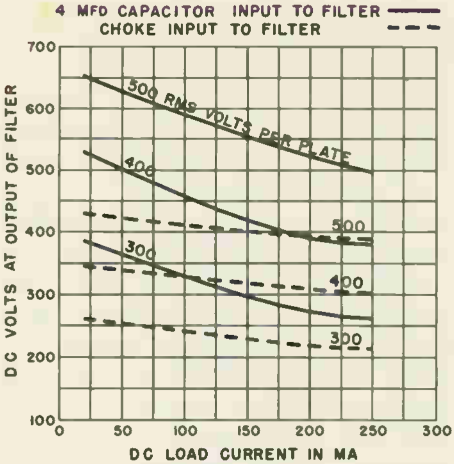


Fig. 21. How the output voltage of a filter network varies for different values of load currents and input voltages. The solid curves are for a capacitor-input filter; the dashed curves for a choke-input filter.

SUMMARY

In this section we have covered some of the more important types of filter networks you are likely to encounter in your career as an electronics technician. You have seen that these networks vary from comparatively simple filters consisting only of a capacitor up to networks containing two chokes and two filter capacitors.

Simpler types of filter networks can be used where the current drain is low and where the filtering does not have to be very good. The more elaborate filters are used in power supplies having a high current drain and in cases where good filtering is required to supply pure dc at the power supply output.

The circuits in this section of the lesson are shown with a half-wave rectifier. The same filter circuits are also used with full-wave rectifiers.

Study the circuits shown in this section. You should be familiar with each of these circuits. You might try drawing these circuits several times to get familiar with the circuit arrangement. Copy them from the book the first few times you try drawing them and then try to reproduce the circuit from memory. This will help you remember what the circuits look like, and if you can remember what they look like, the chances are that you'll be able to remember how they work.

SELF-TEST QUESTIONS

(i) To what value may the input filter capacitor charge in a simple capacitor filter such as shown in Fig. 11?

(j) In what type of application may a simple capacitor-type filter (such as shown in Fig. 11) be used?

(k) What advantage does an R-C pi-type filter (such as shown in Fig. 13) have over the simple capacitor-type filter shown in Fig. 11?

(l) What is the disadvantage of the R-C pi-type filter, and how can this disadvantage be overcome?

(m) What is the purpose of the resistor R1 in the power supply shown in Fig. 17?

(n) Why are choke-input filters used with mercury-vapor rectifier tubes?

(o) What advantage does a choke-input filter have over a capacitor-input filter?

(p) What is a swinging choke?

Typical Power Supplies

The two main sections of the power supply are the rectifier and the filter section. Now that we have discussed both these sections, let's examine some typical power supplies and see what they look like. First we'll look at some power supplies using tube-type rectifiers which might be found in equipment using vacuum tubes. Remember that in these circuits the rectifier operates in the same way as a selenium or silicon rectifier would operate. Insofar as the rectification action is concerned, it makes little difference whether a tube, a selenium rectifier, or a silicon rectifier is used. In the circuits where tubes are used, we will in some cases show how the tube heaters are connected. After looking at a few tube circuits we will look at some typical power supplies using silicon rectifiers and then at a more complex regulated supply.

UNIVERSAL AC-DC POWER SUPPLIES

The universal ac-dc power supply is so called because it can be operated from either an ac or a dc power line. When these power supplies were first used in radio receivers, manufacturers played this feature up, but actually this type of power supply is used to keep costs at a minimum.

The circuit of an ac-dc power supply for a 5-tube radio receiver is shown in Fig. 22. This supply not only supplies the dc voltages required by the plates and screens of the tubes in the receiver but also contains the heater supply for the heaters of the various tubes. First, let us look at the heater supply. Notice that the heaters of the various tubes are connected in series and that they are operated on ac. In this type of power supply you will always find that the heater of the rectifier tube, in this case a 35W4 tube, is connected directly to one side of the power line. The heater of this particular tube is tapped, and a pilot light is connected in parallel with one part of the heater. The pilot light will light when the set is turned on. The tube is designed to be operated with the pilot light connected in parallel with part of the heater, so if in servicing a receiver using a tube with a pilot light tap you should find that the pilot light is burned out, it is a good idea to replace it to keep the heater current in the rectifier tube within its rated value.

Next to the 35W4 rectifier tube in the heater circuit (or heater string, as it is usually called) you will always find the high-voltage power output tube. By a high-voltage tube we mean a tube that requires a high heater voltage. In the diagram we have shown, the tube is a 50B5 tube. As the 50 suggests, a heater voltage of 50 volts is required to operate the tube.

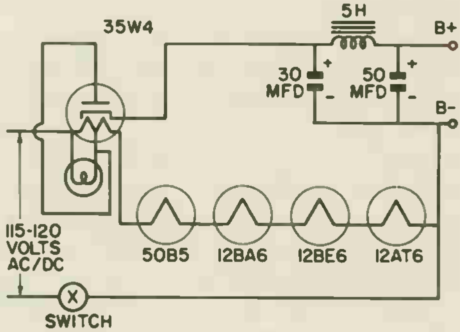


Fig. 22. A typical ac-dc power supply.

Next to the output tube you will usually find the i-f tube. In this diagram, the i-f tube is a 12BA6 tube. It is a tube with a 12.6-volt heater.

Next to the i-f tube you will find the converter tube. The 12BE6 tube is a converter tube designed so that the single tube performs two tasks: one part of the tube is the oscillator and the other part is the mixer.

The last tube in the heater string is always the first audio tube. In this circuit it is a 12AT6 tube. This tube is placed at the end of the heater string nearest B- to keep hum at a minimum, because hum picked up in the first audio tube will be amplified by the entire audio system and could be objectionable. The 12AT6 tube is a dual-diode-triode tube: it has two diodes and a triode inside the same glass envelope. As the 12 preceding the type designation suggests, the tube requires a heater voltage of about 12 volts--to be exact, 12.6 volts.

Now if we examine the rectifier circuit you will see that the plate is connected to a center tap on the rectifier heater. This means that the B supply current flowing through the plate of the tube must flow either through part of the rectifier heater or through the pilot light back to one side of the power line. The purpose of connecting the plate this way is to provide some protection in the event of a short in the receiver. If there is a short in the receiver, the rectifier tube will begin passing excessive current. If this current becomes too high, the pilot light and half the rectifier heater will burn out, opening the circuit and protecting the receiver and the house wiring.

The rest of the power supply is similar to the circuit you studied already. The filter network consists of a capacitor-input filter; the input filter capacitor is the 30-mfd capacitor, and the output filter capacitor is the 50-mfd capacitor. Usually these two capacitors are in a single container that has only three leads brought out from it. Since the negative leads are both connected to B-, a common negative lead and two separate positive leads are usually used. The capacitor is usually a tubular type of capacitor with a paper cover impregnated with wax. The capacitor is mounted on the receiver chassis by means of a mounting strap, or in a receiver with a printed circuit board, the leads are soldered directly to the circuit board. Often this is the only kind of mechanical mounting used.

Capacitors in this type of receiver are usually rated at 150 volts. The normal output voltage of this type of power supply under load is usually somewhere between 90 and 105 volts.

Not all universal ac-dc power supplies use a filter choke. In the circuit shown in Fig. 22, the plates and screens of all of the tubes are operated from the B+ output of this power supply. However, sometimes you will find a filter resistor used in the power supply in place of the filter choke. When this is done, the plate of the output tube is usually connected directly to the cathode of the rectifier tube so that the current drawn by the output tube plate will not flow through the filter resistor. This tube usually draws more current than all the other tubes in the receiver combined.

The output tube is normally a beam tube or a pentode tube, and the plate current depends very little on plate voltage. Therefore if there is some hum voltage applied to the plate of the tube, it usually does not cause any plate current variation and hence is not heard as hum in the output from the loudspeaker. The screen voltage for the output tube and the plate and screen voltages for the remaining tubes are obtained at the output of the power supply where the additional filtering obtained from the filter resistor and the output filter capacitor will reduce the ripple from the rectifier to a low value.

In some of the later table-model radio receivers only four tubes are used; a selenium rectifier or a silicon rectifier is used in place of a rectifier tube because either of these has a considerably longer life. In receivers of this type, tubes with higher heater voltage requirements are often used, so that the sum of the heater voltages required by the four tubes adds up to about 120 volts. If the heater voltage required by the tubes is less than the line voltage, a voltage-dropping resistor can be placed in series with the heaters to use up the leftover voltage. This will cause the total voltage drop across the series-dropping resistor and the tubes to be equal to the line voltage, thus allowing each of the tubes to have its required heater voltage.

Of course, in any heater string where tube heaters are connected in series, all the tubes must have the same heater current rating. Tubes designed for this type of service have what is called a controlled warm-up. This means that the tube heaters are made so that they all reach operating temperature at the same time. This prevents one tube from warming up too quickly and from having too high a voltage across its heater.

The voltage across the tube heater will depend upon its resistance, which changes with temperature. As long as all the tubes warm up at the same rate, their resistances change at the same rate. In this way, a high voltage across any of the tubes is avoided.

Before leaving our study of this type of power supply, it is worthwhile to consider what will happen if a tube such as the 12BA6 tube in Fig. 22 develops a cathode-to-heater shortage. The cathode will usually be connected either to B- directly or to B- through a low-value resistor; this will effectively short out the heaters of the 12BE6 and 12AT6 tubes and, as a result, these tubes will not light.

If you're called upon to service a receiver of this type and you find that one or more of the tubes is not heating, look for a cathode-to-heater short in either the tube which doesn't light that is highest up on the string, or the tube preceding it. If none of the tubes lights this is an indication, since the tubes are connected in series, that the series string is open -- chances are that the heater of one of the tubes has burned out.

A TYPICAL FULL-WAVE POWER SUPPLY

Fig. 23 illustrates a typical power supply using a power transformer and full-wave rectifier. This type is found in many radio and TV receivers and transmitting equipment, as well as in the equipment used in industrial electronics.

Notice that the power transformer has a primary winding to which the 115-volt ac power line is connected; a high-voltage secondary with its center tap grounded, and the two end leads connected to the plates of the rectifier tube; and the two filament

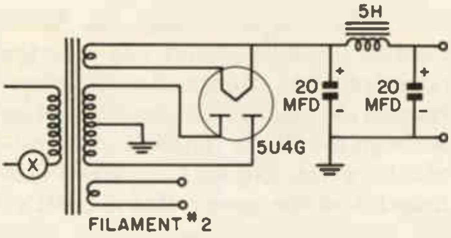


Fig. 23. A typical power supply using a transformer and full-wave rectifier.

windings. One filament winding is used to supply the heater voltage required by the rectifier tube, and one filament winding is used to supply the heater voltage for all the other tubes in the set.

In this power supply we have shown a 5U4G rectifier tube, which requires a heater voltage of 5 volts and a heater current of 3 amps. Thus the rectifier filament winding on the transformer must be capable of supplying 5 volts at a current of 3 amperes.

The winding marked filament number two is used to supply the heater voltage required by all the other tubes in the set. You will notice that this winding is called the filament winding, not the heater winding. This is a carry-over from the old days of radio when most of the tubes were filament-type tubes and few had a separate cathode and heater.

In an electronic device using this type of power supply, the heaters of the tubes (with the exception of the rectifier tube) are connected in parallel. Since the tubes are connected in parallel, you will find that all of the tubes are designed to operate from the same heater voltage. In most cases equipment using this type of power supply has tubes with heater voltage ratings of 6.3 volts. The filament winding must be capable of supplying the heater current required by all of the tubes. The tubes may require different heater currents since they are connected in parallel. You can determine the total current the winding must supply by looking up the heater current required by the individual tubes in a tube manual and adding these figures together.

Again, the filter network used in this power supply is a capacitor-input type. Notice that the capacitors have a lower capacity than those used in the circuit shown in Fig. 22. It is not as necessary to use large capacitors in a full-wave rectifier type of power supply as it is in a half-wave rectifier to obtain the same amount of filtering. Of course, in some power supplies in which it is essential to keep the hum voltage very low, you will find larger filter capacitors. You may also find a two-section filter using an additional choke and a third filter capacitor.

In a power supply such as the one shown in Fig. 23, the two capacitors will probably be mounted in a single container. If a cardboard tubular type is used, one common negative lead and two separate positive leads will be brought out of the container. The same color leads will probably be used for the two positive sections since they both have the same capacity. In some pieces of equipment a metal can-type capacitor might be used -- with this type the can is usually the negative terminal. Mounting the capacitor on the metal chassis automatically makes the connection between the negative terminal and the chassis. The positive leads are brought out of two separate terminals.

DUAL-VOLTAGE POWER SUPPLIES

A diagram of a typical dual-voltage power supply is shown in Fig. 24. This diagram is actually quite similar to the diagram of circuits used in a modern color-TV receiver. Notice that a voltage that is negative with respect to ground is developed by the diode D1 and its associated circuitry. Diodes D2 and D3 are used in a half-wave voltage doubler circuit.

In this circuit when terminal A is negative with respect to terminal B, current flows from A through R1 and D1 to charge capacitor C1. A simple pi-type R-C filter is used in this section of the power supply because the current requirements are low and there will be very little voltage drop across R2. At the same time with the two capacitors, small value capacitors can be used and adequate filtering obtained.

When terminal A is negative, current also flows through the thermistor R3 through R4 and into the negative plate of capacitor C3. Electrons flow out of the positive plate through the diode D2 back to terminal B of the power line. When terminal A of the power line is positive and terminal B is negative the voltage will be placed in series with the voltage built up across C3, so that capacitor C4 is charged to a value approaching twice the peak line voltage through diode D3. This provides an output voltage which is positive with respect to ground and approximately equal to twice the ac line voltage.

The thermistor R3 is put in the circuit to prevent high current surges through the silicon diodes D2 and D3 when the equipment is first turned on. R4 is used in the circuit to provide further protection. If the equipment is turned on and operating for some time, the resistance of the thermistor will drop to a low value. If the equipment is turned off for a

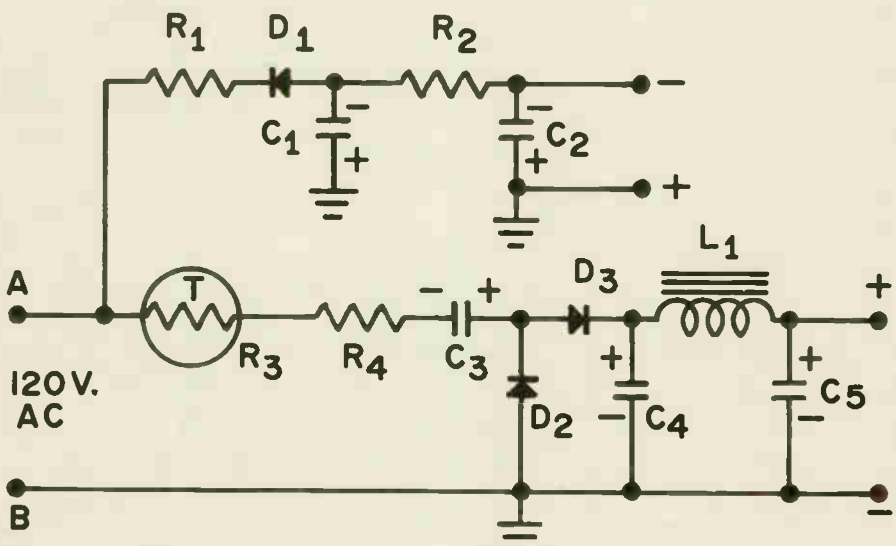


Fig. 24. A typical dual-voltage power supply.

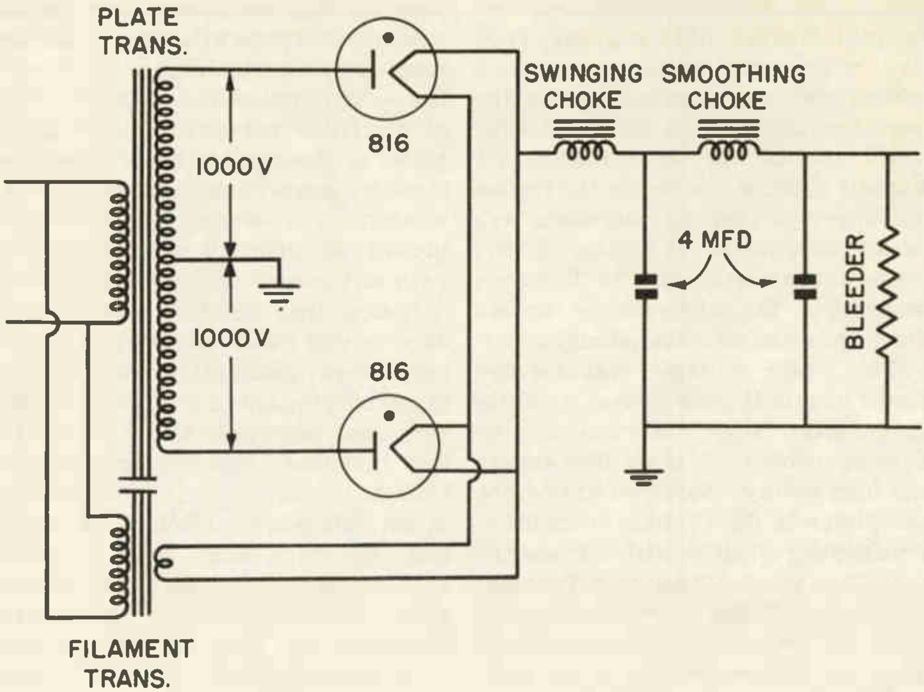


Fig. 25. A typical transmitter power supply.

few seconds and then turned back on, the charging current through the diodes could be quite high and damage them. R4 is put in the circuit; its value does not change and it limits the current through the diodes to prevent their damage under this circumstance.

The power supply shown in Fig. 24 is typical of the power supply you are likely to run into in both monochrome and color television receivers. A voltage-divider network may be used across the output of either supply to provide different value voltages. This power supply is comparatively inexpensive; it eliminates the need of a power transformer and with modern silicon diodes is comparatively trouble-free.

HIGH-VOLTAGE POWER SUPPLIES

Another power supply is shown in Fig. 25. This is the type of power supply you will find in transmitting equipment or in other equipment where high operating voltages are required. Notice that in this power supply, full-wave rectification is used. Also notice that it has two separate rectifier tubes. Separate rectifier tubes rather than a single tube with two plates are used in high-voltage supplies because the voltages are so high that they would simply arc across inside the tube. The type 816 tubes shown in this supply are mercury-vapor tubes designed for use in power supplies where the operating voltages are not too high and where the current requirements are not too great. They are often used in power supplies where the ac input voltage to the rectifier is between 1000 and 2000 volts, and the current drain does not exceed 250 ma. As far as the transmitting-type power supplies are concerned, an output voltage of 1000 volts across each half of the secondary of the high-voltage transformer is not considered high.

The high-voltage transformer found in this type of power supply is frequently called a plate transformer because it is used to supply the high voltage required to operate the plates of the various tubes in the transmitter. The rectifier tubes are operated from a separate filament transformer. In a power supply using type 816 tubes, the filament voltage required by these tubes is 2.5 volts and the current required is 2 amps. Therefore, this filament transformer must be capable of supplying a total of 4 amps at a voltage of 2.5 volts. The filament transformer must have good insulation between the secondary winding, the transformer core, and the primary winding; otherwise the high voltage will arc through the insulation either to the primary winding or to the transformer core.

At the input of the power supply filter network is a swinging choke. This is common practice in transmitter power supplies because it improves the voltage regulation, and also because it affords additional protection for the mercury-vapor rectifier tubes.

The second choke in the power supply is called a smoothing choke. This is the same type of choke shown in the power supplies in Figs. 22 and 23. It is called a smoothing choke in transmitting equipment because its primary purpose is to smooth out the ripple and also to distinguish it from the swinging choke used at the input of the filter network. A smoothing choke is designed to keep its inductance as nearly constant as possible, whereas a swinging choke is designed so that its inductance will vary as the current through it varies.

Notice that the filter capacitors used in this power supply are 4-mfd capacitors. Each of these capacitors is usually mounted in its own separate container. Occasionally you will find two small high-voltage capacitors in the same container, but this is not common practice. Also notice that the capacitors have a much smaller capacity than those in the two power supplies we discussed previously. These capacitors are oil-filled capacitors and it is quite costly to make this type of capacitor with a large capacity. On the other hand, capacitors used in circuits like those in Fig. 22, Fig. 23, and Fig. 24 are electrolytic capacitors and very large capacities can be obtained at a very moderate cost.

Effective filtering is obtained with the smaller-size capacitors in this supply, because two filter chokes are used. The inductance and hence the inductive reactance of these chokes are usually somewhat higher than that of filter chokes used in lower voltage equipment. The choice of using either large chokes or large filter capacitors is simply one of cost. At low voltages it is more economical to use large capacitors and low-inductance chokes, but at high voltages it is more economical to use high-inductance chokes and low capacities. The net result is the same as far as the filtering action is concerned. Another component that you will find in transmitting-type power supplies is a bleeder. A bleeder is a resistor connected across the power supply output. The bleeder serves two purposes: it improves the voltage regulation and is also used for safety.

A bleeder resistor connected across the power supply keeps the minimum current at a reasonable value. If the current drawn by the load connected to this power supply were to drop to zero, the two filter capacitors would be charged up to a value equal to the peak voltage across half of the secondary of the plate transformer. If the transformer had an rms voltage of 1000 volts across each half of the secondary, this would mean that the capacitors would charge up to a voltage of about 1400 volts. This may be high enough to destroy the capacitors. Also, the chokes have a definite maximum voltage that can be applied to them. If the voltage goes too high, the insulation between the choke winding and the core may break down. This will destroy the chokes. If either of the chokes or capacitors shorts, the rectifier tubes will pass such a high current that they may be ruined. There is also the danger of burning out the plate transformer. Furthermore, if the voltage reaches too high a value, the rectifier tubes may arc over internally. A bleeder connected across the power supply output can eliminate this danger. With the bleeder across the output, if the equipment current drops to or almost to zero, the bleeder current will continue to flow. If the output voltage starts to rise, the bleeder current will increase because the current flowing through any resistor increases if the voltage across it increases. The bleeder current will keep the voltage from climbing to an unsafe value.

Another important reason for using the bleeder is that an oil-filled capacitor such as those found in this type of power supply can hold a charge for a long time. If the two 4-mfd capacitors used in the supply were charged up to a voltage of 1000 volts or more and a technician servicing the equipment accidentally touched one of these capacitors, he could receive a very dangerous shock. Under certain conditions it could be fatal. This danger can be greatly reduced by connecting a bleeder across the power supply so that when the equipment is turned off the capacitors are discharged through the bleeder.

You may have occasion to work on high-voltage power supplies at some future date. Remember that a bleeder is connected across the power supply for safety as well as to improve the voltage regulation. Therefore if the bleeder in a power supply burns out, it should be replaced. However, never rely on a bleeder to discharge high-voltage filter capacitors. If you have to work on a high-voltage power supply, your first step should be to remove all voltages from the supply. To do this, turn the power supply off; if there are fuses in it, remove the fuses so that no one can accidentally turn it on; disconnect it completely from the source if you can. Often it is not possible or convenient to completely disconnect the equipment from the voltage source but if it is shut off and any fuse in the circuit is removed, it should be safe. Next, before you start to work on the supply, discharge all filter capacitors in the power supply. The capacitors should be discharged with a heavy metal rod that has a good insulated handle so that you will not come in contact with the metal rod. Use the metal rod to short together the terminals of the capacitor to discharge it. Touch the grounded terminal of the capacitor first and slide the rod over to touch the other terminal. Do this several times to be sure the charge is completely removed. After the capacitors have been discharged, the power supply should be safe to work on.

Keep this point in mind: high voltage capacitors, or for that matter any large capacitors, should be discharged before you start to work on a piece of equipment. Many technicians fail to do this. There are some technicians who can tell about the terrific shock they received when they failed to discharge a filter capacitor. There are others that did not survive the experience to tell about it.

A TRANSISTOR-REGULATED POWER SUPPLY

A regulated power supply employing transistor voltage regulators is shown in Fig. 26. This power supply is used in a TV receiver that is designed for operation from the power line and also from a 12-volt dc source. When the power plug is plugged into a 120-volt line and the switch is turned on, the receiver will operate from the power line. When the power plug is disconnected and switch S1 is closed, the receiver can be operated from a 12-volt battery.

The operation of the power supply from the power line is comparatively simple. Two diodes, D1 and D2, are used in a full-wave rectifier circuit. When the transformer T1, which is a step-down transformer, has a polarity such that the end of the secondary connected to D1 is negative, current will flow through the diode D1 to ground and into the negative plate of C3. Electrons flow out of the positive plate of C3 to the center tap of the power transformer which is positive with respect to the end connected to D1. During the next half-cycle, when the end of the secondary connected to D2 is negative, current will flow through D2 to ground, into the negative plate of C3,

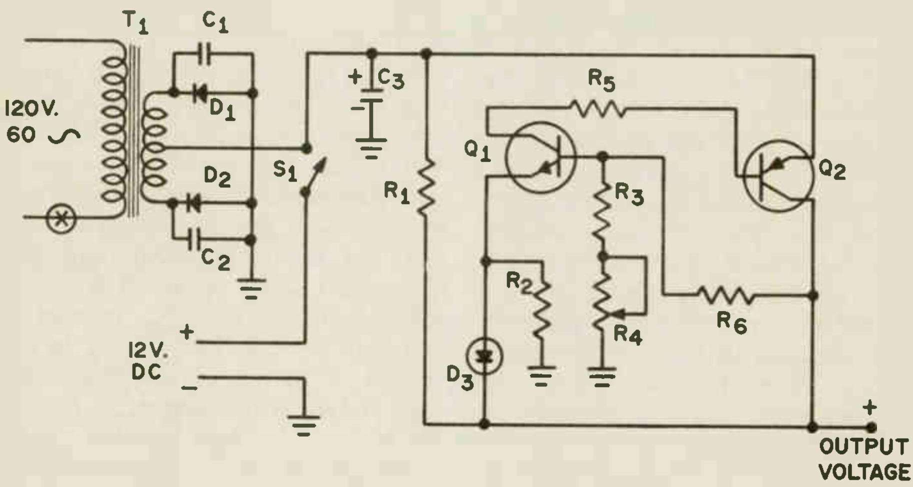


Fig. 26. A transistor-regulated power supply.

out of the positive plate of C3 and back to the center tap of the secondary winding on the power transformer.

The remainder of the components used for the power supply are used for the purpose of regulating the voltage. In other words, the power supply voltage is maintained constant at approximately 12 volts regardless of the load drawn from the supply. The transistor Q2 is a PNP transistor that is used as a series voltage regulator. Notice that the emitter of this transistor connects directly to the positive side of C3. You can consider this transistor as working more or less as a variable resistor: if the output voltage tends to rise, the resistance increases and if the voltage tends to fall, the resistance decreases.

The effective resistance of Q2 is varied by varying the forward bias across the emitter-base junction. Notice the zener diode D3. This diode is connected in series with R2. The zener has a constant voltage of 6.3 volts across it. Therefore, the voltage drop across R2 will be equal to the output voltage minus 6.3 volts. This is the emitter voltage applied to Q1. The base voltage is determined by the voltage division occurring between R4, R3 and R6. R4 is adjustable so that the output voltage can be adjusted to 12 volts. Under these circumstances a certain current will flow through Q1 and through R5 and this will set the forward bias on Q2. If the output voltage tends to rise, the base voltage on Q1 will rise but by an amount less than the emitter voltage. The divider network consisting of R4, R3 and R6 will prevent the base from rising the full amount of the output voltage rise. On the other hand, the voltage across the zener D3 remains constant so that the voltage across R2 will reflect the entire output voltage rise. This will reduce the forward bias on Q1 which, in turn, will reduce the emitter-collector current. The reduction in the emitter-collector current will reduce the voltage drop across R5 which, in turn, will reduce the forward bias on Q2; this has the effect of increasing its resistance. The increased resistance tends to keep the output voltage from increasing.

If the output voltage decreases, the opposite happens. The base voltage on Q1 falls, as does the emitter voltage. However, the emitter voltage falls more than the base voltage, so the forward bias is increased. This increases the emitter-to-collector current through Q1 which increases the forward bias on Q2. This has the effect of reducing the resistance on Q2 and tends to keep the output voltage from falling.

This type of power supply is one of the more complex power supplies that you are likely to encounter in electronic equipment. The voltage regulation is required in order to keep the voltage reasonably constant on the various transistors used on the TV receiver. In most cases, such precise voltage regulation is not required in entertainment-type equipment.

VOLTAGE DIVIDERS

We mentioned previously that more than one operating voltage is sometimes needed in the various stages of a piece of electronic equipment. Rather than use a separate supply for each voltage needed, the usual procedure is to use a single supply designed to give the highest voltage needed, and then obtain the lower voltages required by means of a voltage divider connected across the power supply output. A typical voltage divider is shown in Fig. 27.

In this voltage divider, R1 and R2 are voltage-dropping resistors; they drop the voltage from 300 volts to the required voltages of 200 and 100 volts. R3 is a bleeder used to stabilize the voltages at points B and C. With this type of network, terminal D is the ground or common terminal. Between D and C there is a voltage of 100 volts; terminal C is positive with respect to terminal D. Between D and B there is a voltage of 200 volts and terminal B is positive with respect to terminal D. Finally, between terminals D and A there is the full power supply output voltage of 300 volts, and of course terminal A is positive with respect to terminal D.

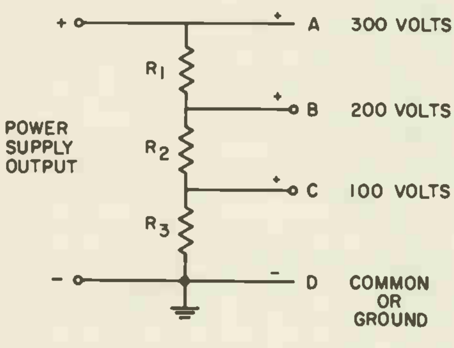


Fig. 27. A voltage-divider network.

The current flowing through R3 is called the bleeder current. It remains fairly constant and is determined primarily by the sizes of R1, R2 and R3. Usually, the size of R3 is chosen so that the bleeder current will be at least as great as the current drawn by the stages connected to terminals C and B. Choosing a value of R3 that will result in a reasonable bleeder current helps maintain good voltage regulation at terminals C and B.

The current flowing through R2 will be made up of the bleeder current plus the current drawn by the stages connected to terminal C. If this current varies, the voltage drop across R2 and hence the voltage at terminal C will vary. However, the bleeder current will remain essentially constant so that if a sizable percentage of the current flowing through R2 is bleeder current, variations in the current drawn by the stages connected to terminal C do not cause too much variation in the voltage drop across R2.

The current flowing through R1 is made up of the bleeder current plus the current drawn by the stages connected to terminals C and B. Again if the bleeder current through R1 represents a sizable part of the total current flow through R1, variations in the current drawn by the stages connected to terminals C and B do not cause too great a variation in the voltage drop across R1 so the voltage at terminal B will remain reasonably constant.

Bleeders are not used in modern midget radio receivers, but you will find them in many of the older sets. They are frequently used in TV receivers, in the low-voltage power supplies in transmitting equipment, and in industrial electronic equipment.

Sometimes one section of a tapped resistor will burn out. Often you can repair the equipment simply by connecting a resistor having the correct resistance and a suitable wattage rating across the defective section. Of course, if separate resistors are used in the voltage divider you can simply replace any defective one. If you do shunt a burned out section of a tapped resistor in a radio receiver and find the equipment is noisy after you have made this repair, the defective section may be making contact intermittently and creating the noise. Of course in this case you must replace the entire unit either with separate resistors connected in series or with a tapped resistor like the original one.

VIBRATOR-TYPE SUPPLIES

The radios installed in automobiles for years used a power supply known as a vibrator type of power supply. A schematic diagram of this type of power supply is shown in Fig. 28. The heart of this type of power supply is the vibrator, which is used to change the dc from the automobile storage battery to a pulsating current in the primary winding of the power transformer.

The vibrator consists of an electromagnet L, and a reed (R-K) placed between two sets of contacts. In the circuit shown in Fig. 28, when the switch is turned to the ON position, current will flow from the negative terminal of the battery through the switch and through the reed towards terminal M. Here it will flow from the reed to contact M, to coil L, through coil L back to the positive side of the battery. The current flowing through the coil creates a magnetic field. This magnetic field attracts the end of the reed K, pulling the reed over toward L and contact N. When the reed makes contact with terminal N, current flows through the upper half of the transformer primary winding. It flows from the top of the winding to the center tap, building up a magnetic field.

At the same instant that the reed is making its contact with terminal N, it will break its contact with terminal M so that the electromagnet will no longer be energized and the field about it will collapse. The reed is made of a spring type material so that it springs back until it makes contact with both terminal M and terminal O. At the instant contact is made with terminal O, current flows through the lower half of the primary winding of the transformer, flowing from the bottom of the winding towards the center tap. The current is flowing through the primary winding in the opposite direction to the direction in which it was flowing through the upper half of the trans-

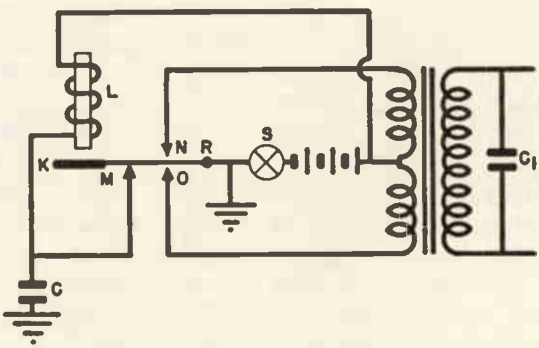


Fig. 28. A vibrator circuit.

former winding. Thus we have a field built up in one direction and then in the opposite direction. At the same time the fact that the reed makes contact with terminal M will once again complete the circuit through the electromagnet so the reed will swing over to the magnet again, making contact with terminal N. As you can see this action causes the reed to vibrate back and forth between terminals N and O. Thus we have a field built up in the primary first in one direction and then in the opposite direction. Building up this field, collapsing it, and then building up a reverse field and collapsing it, means that we have a continually changing magnetic field cutting the secondary of the transformer. By putting a large enough number of turns on the secondary, we can obtain whatever voltage we may require for the operation of the receiver.

A complete vibrator-type power supply is shown in Fig. 29. Notice that the secondary of the vibrator transformer is center tapped and that a full-wave rectifier is used. The capacitor C_2 is called a buffer capacitor. This is a high voltage paper capacitor that is used to keep sharp noise pulses out of the power supply. The size is usually quite critical and if it is necessary to replace the buffer capacitor in a receiver using this type of power supply, you should use a capacitor having the same capacity as the original.

In a vibrator-type power supply there is considerable sparking as the reed vibrates back and forth. This sets up a radio-frequency type of interference which could get through the rectifier and cause considerable interference in the receiver. In the power supply shown in Fig. 29, the choke L and the capacitor C are called hash suppressors. This rf interference or noise is called hash; the choke and the capacitor are put in the power supply in order to keep as much as possible of this hash or noise out of the power supply output. Capacitor C acts like a short circuit to these radio frequency pulses, and choke L acts like a very high impedance to them. Thus L and C form a voltage-divider network, with most of the voltage appearing across the high impedance L and little or no voltage across the low impedance C.

Vibrator-type power supplies were used in almost all automobile receivers in automobiles using 6-volt ignition systems. However, in newer cars, a 12-volt ignition system is used. The first receivers made for these cars also used vibrator type supplies, but tube manufacturers designed special tubes that will operate with plate and screen voltages as low as 12 volts. These 12-volt tubes were used in automobile receivers for a few years, but they too were replaced by transistors. Since transistors operate from low voltages, the vibrator type of power supply is no longer needed. These supplies were not only costly, but in addition they were one of the most troublesome sections of the automobile receiver.

SUMMARY

The power supplies we have shown in this section of this lesson are typical of the various types of power supplies you are likely to encounter as an electronics technician. You will find many variations of these circuits, but these are the basic circuits. Spend some time studying

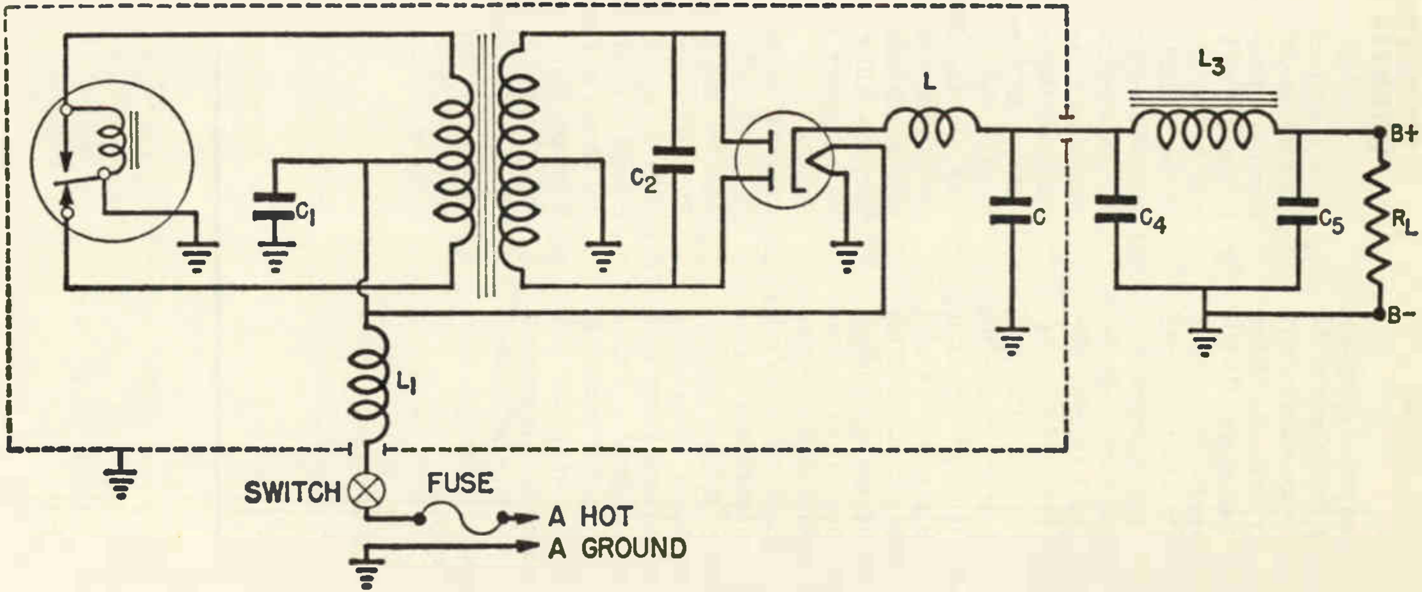


Fig. 29. A complete vibrator power supply.

these circuits so you will know what they look like.

In servicing these power supplies, keep in mind that the values of the various components used are generally not extremely critical. Manufacturers usually use as small a filter capacitor as they can in a circuit. It is more economical to use a small capacitor than to use a large one. Therefore, if you have an ac-dc receiver that uses a 20-40 mfd, 150-volt capacitor and you find it necessary to replace the filter capacitor, there is no reason why you could not use a 30-50 mfd, 150-volt capacitor in its place. Some variation of part values can be made without affecting the performance of the equipment.

It is also possible to use capacitors having a higher voltage rating than those used originally. This will simply provide an additional margin of safety. You may find that you cannot do this in large transmitter power supplies because there isn't room to mount a capacitor with a higher voltage rating, but usually in radio, TV, and most small pieces of industrial electronic equipment there is enough room to put in a capacitor of slightly larger physical size than the original.

SELF-TEST QUESTIONS

(q) Draw a schematic diagram of a typical ac-dc power supply found in a five-tube radio.

(r) In the circuit shown in Fig. 22, why is the 12AT6 tube placed at the B- end of the string?

(s) If in an ac-dc receiver in which the heaters of the tubes are connected in series and none of the tubes lights what would you suspect the cause of the trouble to be?

(t) Why are the filter capacitors used in the power supply shown in Fig. 23 smaller in capacity than the filter capacitors used in the power supply shown in Fig. 22?

(u) What is the purpose of R3 in Fig. 24?

(v) What is the purpose of R4 in Fig. 24?

(w) What type of voltage-doubler circuit is used in the power supply shown in Fig. 24?

(x) What purpose does the transistor Q2 serve in the power supply shown in Fig. 26?

(y) What is the purpose of the diode D3 in Fig. 26?

ANSWERS TO SELF-TEST QUESTIONS

(a) There will be 60 current pulses per second through the load.

(b) The disadvantage of the half-wave rectifier circuit is that current flows through the rectifier during one half-cycle and not during the other half-cycle. As a result, the output is somewhat difficult to filter and smooth out to pure dc.

(c) The diode D1 in the circuit shown in Fig. 6 is used to charge the capacitor C1 during one half of each cycle. Capacitor C1 is charged so that during the next half-cycle the voltage across it will be in series with the line voltage. This will place a voltage equal to twice the line voltage across the load and diode D2. Since diode D2 has a very low resistance when it is conducting, the voltage across the load is twice what it would be without the combination of C1 and D1 in the circuit: hence, the circuit is called a voltage-doubler circuit.

(d) The circuit is called a full-wave rectifier circuit because a current pulse flows through the load during each half-cycle. In other words, if the rectifier circuit is operating from a 60-cycle power line there will be 120 current pulses through the load (one for each half-cycle).

(e) The disadvantage of this circuit is that the high voltage winding on the power transformer must be center-tapped. This means that the high-voltage winding on the transformer must have twice the number of turns required to get the desired output voltage across the load. This circuit requires a rather expensive power transformer.

(f) The advantage of the bridge rectifier circuit is that there is a saving in the power transformer cost over that of a transformer that has a tapped high-voltage secondary winding; also, the circuit is capable of good voltage regulation.

(g) The voltage-doubler circuit shown in Fig. 9 is a full-wave voltage doubler. That is, there will be 120 current pulses per second in the output of the voltage doubling capacitor network consisting of C1 and C2. The circuit shown in Fig. 6 is a half-wave voltage-doubler circuit and there will be only 60 current pulses through the load in this circuit. It will be somewhat easier to filter and smooth the output voltage in the circuit shown in Fig. 9 than it will be in the circuit shown in Fig. 6.

(h) A full-wave doubler circuit requires a less expensive power transformer for a given load voltage than the bridge rectifier circuit requires. Also, the voltage-doubler circuit requires only two rectifiers whereas the bridge-rectifier circuit requires four rectifiers. The disadvantage of the full-wave voltage-doubler circuit is that it does not have as good voltage regulation as the bridge-rectifier circuit.

(i) The capacitor in a simple filter circuit such as shown in Fig. 11 may charge up to a value equal to the peak value of the ac input voltage. In the case of a power supply operating from a 120-volt line this is equal to approximately 1.4 times 120 volts.

(j) A simple capacitor-type filter may be used in applications where the current drain is low. With a low current drain the capacitor discharges very little between cycles so that the voltage across the capacitor, and hence the voltage across the load, remains essentially constant.

(k) The R-C pi-type filter is capable of better hum elimination than a simple capacitor-type filter. This type of filter is particularly desirable where the current drain is high enough to discharge the capacitor appreciably between charging cycles in a simple capacitor-type filter.

(l) The disadvantage of the R-C pi-type filter is that there is considerable voltage drop across the filter resistor. This problem can be overcome by using a filter choke such as in the L-C type filter shown in Fig. 15. A filter choke will offer a high opposition to any ac and thus effectively reduce the ac, while at the same time offering a low resistance to the passage of dc through it.

(m) R1 in the power supply shown in Fig. 17 is used to limit the current through the silicon rectifier when the power supply is first turned on. Without this resistance in the circuit, the charging current through the diode to charge C1 may be so high that the rectifier may be destroyed.

(n) To limit the peak current through the tubes. A mercury vapor rectifier tube is easily damaged by a high peak current. The peak current through the rectifier tube is much lower with a choke-input filter than it is with a capacitor-input filter.

(o) A choke-input filter will provide better regulation than a capacitor-input filter. This means that the voltage across the load will vary less with widely varying currents when the filter is a choke-input filter than it will when the filter is a capacitor-input filter.

(p) A swinging choke is a choke whose inductance changes as the current changes. As the current builds up the choke tends to saturate so that its inductance goes down. This tends to reduce the reactance of the choke and hence helps provide better voltage regulation.

(q) See Fig. 22. If you cannot draw this diagram from memory, copy it from the book. Simply drawing the diagram will help you to become familiar with the circuit and remember it in the future.

(r) The 12AT6 tube is the first audio stage. It is placed at the B- end of the heater string in order to keep hum pick-up in the tube as low as possible. Any hum picked up by this tube will be amplified by the entire audio system.

(s) The chances are that the heater of one of the tubes is open.

(t) The power supply shown in Fig. 23 uses a full-wave rectifier. Therefore there will be 120 pulses per second to charge the filter capacitors. The power supply shown in Fig. 22 is a half-wave power supply and there will be only 60 pulses per second to charge the filter capacitor; therefore, larger capacitors are needed to eliminate hum.

(u) R3 in Fig. 24 is a thermistor. A thermistor has a high cold resistance, but the resistance decreases as the thermistor heats up. The thermistor is used in this power supply to protect the diode rectifiers from high current surges when the power supply is first turned on.

(v) R4 is a fixed resistor that is used to protect the diodes in the event the equipment is turned off and then turned back on almost immediately. Under these conditions the resistance of the thermistor will be too low to provide the required protection for the rectifiers and hence R4, along with the hot resistance of the thermistor, limits the current through the diode rectifiers to a safe value.

(w) A half-wave voltage doubler.

(x) Q2 is in series with the B supply voltage. It operates essentially as a variable resistor and is used to regulate the power supply output voltage and keep it at essentially a constant value.

(y) The diode D3 is the zener diode. It provides a reference voltage so that the voltage variations on the emitter of Q1 will be greater than the voltage variations on the base. Thus, changes in output voltage affect the forward bias of the transistor and hence the conduction through it.

Lesson Questions

Be sure to number your Answer Sheet B201.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What advantage does a silicon rectifier have over a vacuum-tube rectifier?

2. Draw the schematic symbol for a silicon rectifier, and indicate by an arrow the direction in which the current will flow through it.

3. How many pulses per second will you get at the output of a full-wave rectifier that is operated from a 60-cycle power line?

4. What is the purpose of C1 and D1 in the circuit shown in Fig. 6?

5. (a) In the circuit shown in Fig. 6, how many current pulses will there be through the load (60-cycle power line)?
   (b) In the circuit shown in Fig. 9, how many current pulses will there be through the load (60-cycle power line)?

6. In the circuits shown in Fig. 11, what part supplies current to the load when the rectifier is not conducting?

7. Explain the following things in connection with the L-C circuit shown in Fig. 15:
   (a) The action of the choke when ac flows through it.
   (b) The action of the choke when dc flows through it.
   (c) The action of the capacitor when ac flows through it.
   (d) The action of the capacitor when dc flows through it.

8. What type of filter network has better voltage regulation -- the capacitor input or the choke input?

9. If you are servicing a five-tube table model radio that uses a universal ac-dc power supply and you see that two of the tubes are not lighting, where would you look for trouble?

10. How does an increase in output voltage affect Q1 and Q2 in the regulated power supply shown in Fig. 26?

HOW TO START STUDYING

For some people, starting to study is just as hard as getting up in the morning. An alarm clock will work in both cases, so try setting the alarm for a definite study-starting time each day. Start studying promptly and definitely, without sharpening pencils, trimming fingernails or wasting time in other ways.

Beginning is for many people the hardest part of any job they tackle. So formidable does each task appear before starting that they waste the day in dilly-dallying, in day-dreaming, and in wishing they didn't have to do it. The next day and the next after that are the same story. Indecision brings its own delays, making it harder and harder to buckle down to work.

Are you in earnest? Then seize this very minute; begin what you can do or dream you can. Boldness in starting a new lesson is a great moral aid to mastery of that lesson; only begin, and your mind grows alert, eager to keep on working. Begin, and surprisingly soon you will be finished.

J. M. Smith

nri

ACHIEVEMENT THROUGH ELECTRONICS

LOW-FREQUENCY VOLTAGE AND POWER AMPLIFIERS

B202

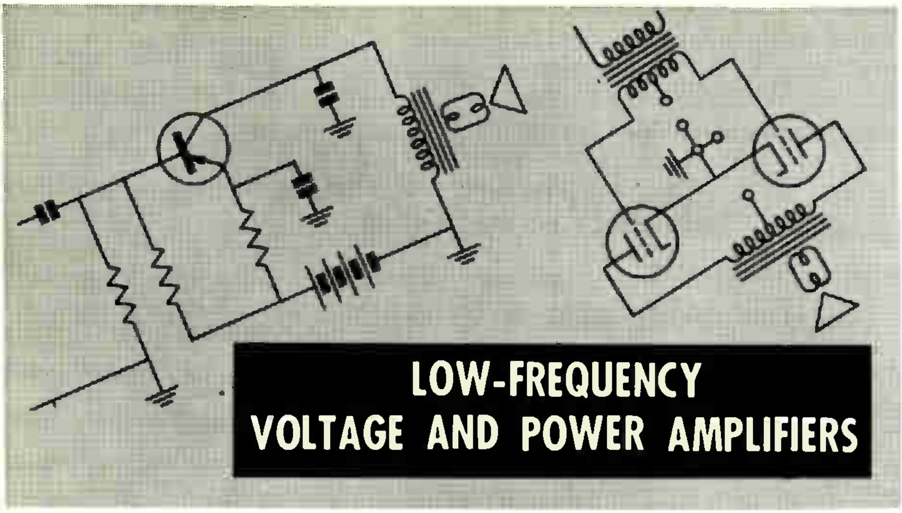
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

LOW-FREQUENCY VOLTAGE AND POWER AMPLIFIERS

B202

STUDY SCHEDULE

JDL5M372 1974 EDITION Litho in U.S.A.



LOW-FREQUENCY VOLTAGE AND POWER AMPLIFIERS

In this lesson and the next two, we will take up the study of amplifiers. For convenience in studying them, we have divided them according to the frequencies they are designed to handle.

You will study low-frequency voltage and power amplifiers in this lesson. These are the kinds used to amplify the sound or audio signal in radio and television receivers, and they are also used in high-fidelity and stereo equipment, in conjunction with amplifying audio signals that are received by means of a radio system from a broadcast station and also with audio signals from a phonograph or tape recorder.

In your next lesson you will study radio frequency amplifiers. These amplifiers are designed to amplify signals received directly from radio and television stations as well as signals within the receiving equipment itself that fall within the radio frequency range. You will study wide-band amplifiers in the third lesson; these are amplifiers that are designed to amplify a wide range of frequencies. For example, a wide-band amplifier might have to amplify signals that have few cycles per second, signals as high as several megacycles per second, and all the frequencies in between these two limits. Wide-band amplifiers of this type are used to amplify the picture signals in television receivers. They are also used in many other special applications.

The ability of tubes and transistors to amplify signals is essentially what makes many of our modern electronic devices possible. Therefore, if you understand how these amplifiers work and how they are put together to perform specific functions, you will be able to analyze the operation of many different types of electronic equipment that you might encounter. Since amplifiers are so important it is worthwhile to spend as much time as necessary on this lesson and on the next two to be sure that you have a complete understanding of them.

The low-frequency amplifiers that you will study in this lesson can be divided into two types: voltage amplifiers and power amplifiers. As you already know, a voltage amplifier is one that is designed to amplify a weak signal voltage and make it stronger. For example, the output voltage from a phonograph pickup in a record player might be only a few thousandths of a volt. This audio signal is too weak to do anything with directly so we feed it to a number of voltage amplifiers to build the amplitude of the voltage up to a usable value. We try to perform this amplification without changing the signal in any way. If we change the signal, we have introduced something called distortion because the amplified signal is no longer the amplified equivalent of the original signal.

Power amplifiers are used to drive the speaker in radio and TV receivers. The speaker requires a certain amount of power in order to cause the cone in the speaker to vibrate back and forth and set the air in front and in back of the speaker into motion. The power required to perform this function is supplied by power amplifiers. The exact amount of power required for the speaker will depend upon the design of the speaker, its size, efficiency and a number of other factors. You will study amplifiers that have a comparatively low power output of perhaps one or two watts and you will also study high-power amplifiers that are capable of putting out 50 or more watts of power.

When we speak of low-frequency amplifiers, we are generally concerned with amplifiers that are designed to amplify signals from about 50 to 100 cycles up to signals of about 10,000 or 20,000 cycles. There are no sharp dividing lines either at the low-frequency or high-frequency end of the range over which the amplifiers are supposed to work. Generally, a low-frequency amplifier is an amplifier that works in the audio range (in other words, within the limits of our hearing). Of course, these amplifiers will amplify signals beyond these frequency limits. The amplifier does not simply stop amplifying at a frequency above the highest frequency it is designed to amplify -- it just doesn't amplify higher frequency signals well. The same is true of low-frequency signals. If an amplifier is designed to amplify signals from about 100 cycles up, it will also amplify a signal having a frequency of 80 cycles, but the chances are that the 80-cycle signal will not receive as much amplification as the 100-cycle signal would.

In this lesson you will learn why the gain of an amplifier falls off as the frequency of the applied signal falls above or below the frequency limits of the amplifier. You will also see that the frequency limits of the amplifier are more or less arbitrarily fixed by design engineers.

In our study of low-frequency amplifiers we will begin with voltage amplifiers. We will study the resistance-capacitance coupled voltage amplifier first because it is the most widely used and hence the most important.

Resistance-Capacitance Coupled Voltage Amplifiers

Resistance-capacitance coupling, usually called R-C coupling or simply resistance coupling, is so named because resistors and capacitors are used to couple the signal from one stage to another.

This type of coupling is widely used between voltage amplifiers and between voltage amplifiers and class A power amplifiers in audio work. It is preferable to transformer coupling in modern equipment because it is more economical and, in addition, usually gives better frequency response. This means it comes closer than transformer coupling does to amplifying equally all signals in the audio range.

Transformer coupling has not been used for many years between voltage amplifiers using vacuum tubes. However, you may find transformer-coupled transistor amplifiers, and there is some advantage to this type of coupling in transistor circuits.

A TYPICAL TUBE CIRCUIT

Fig. 1 shows two typical resistance-capacitance coupled stages. V_1 is the first stage; the output signal from this stage is fed to the second stage (V_2) by means of R-C coupling. The R-C coupling components are R_2 (which is the plate load resistor of V_1), C_3 (which is the coupling capacitor) and R_3 (which is the grid resistance for the second stage, V_2).

Let's review again the operation of these amplifier stages and how the coupling network works. With no signal applied to V_1, a current will flow from the cathode of the tube to the plate because the plate is con-

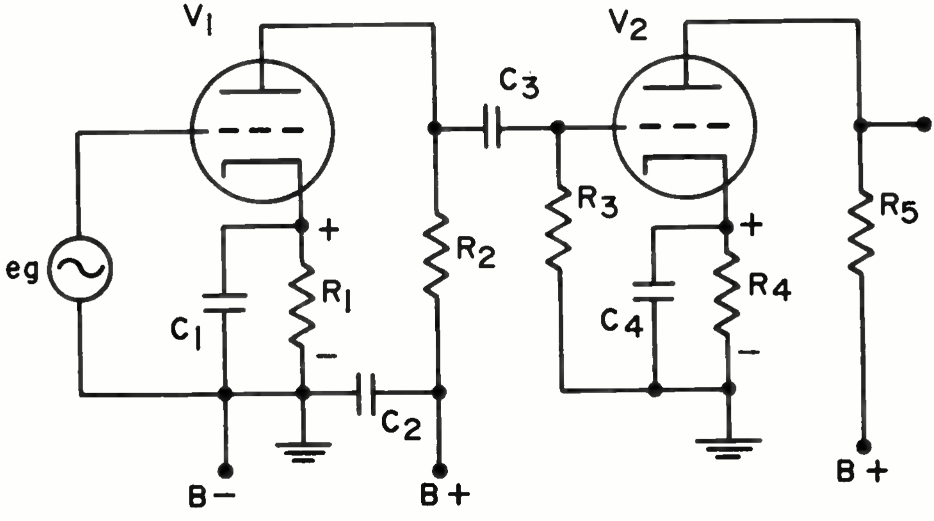


Fig. 1. Typical vacuum tube R-C coupled stage.

nected to the positive side of the B supply through R_2. The cathode is connected to the negative side through R_1. Electrons will leave the negative side of the B supply and flow through R_1 to the cathode. In flowing through R_1 they will set up a voltage drop across the resistor having the polarity shown. This voltage drop will make the cathode positive with respect to ground. The grid of V_1 is returned directly to ground through the generator and since there is no current flowing in the grid circuit there will be no voltage drop across the generator. Therefore the grid is at dc ground potential. This will make the grid negative with respect to cathode. The actual grid-cathode voltage will depend upon the size of R_1 and the current flow through R_1. The current flow through the tube can be determined from a tube manual and the required bias for the tube can be obtained simply by making R_1 large enough to produce the bias voltage needed.

Electrons reaching the cathode will be emitted by the cathode and flow through the tube to the plate. Then the electrons will flow from the plate through R_2 back to the B supply. In flowing through R_2 the electrons will produce a voltage drop across this resistor. The value of R_2 is usually quite high so there will be a substantial voltage drop across the resistor. Thus, although the plate will be positive with respect to ground, this positive voltage on the plate will be less than the positive voltage available at the positive terminal of the B supply.

The current flow through V_2 follows a path equivalent to the current path in V_1. The plate of V_2 is returned to the positive side of the B supply through R_5 so there will be a positive voltage on the plate of V_2. This will cause electrons to flow through R_4, producing a voltage drop across it (having the polarity shown) so that the cathode will be positive with respect to ground and the grid negative with respect to cathode. As before, there will be no current flow in the grid circuit and therefore no current flow through R_3. This means there will be no dc voltage drop across this resistor; consequently, the grid will be at dc ground potential.

When a signal voltage is applied by the generator e_g, the generator voltage will vary the voltage between the grid and cathode of V_1. The generator connects directly to the grid of V_1. The other side of the generator connects to the cathode through C_1. C_1 is chosen so that its reactance will be very low at the frequency of the signal to be handled. Thus we have the input signal from the generator applied directly between the grid and cathode of V_1.

When the polarity of the input signal is such that the grid end of the generator is positive and the other end is negative, the voltage applied between grid and cathode by the generator will subtract from the bias voltage across R_1. This will make the grid less negative with respect to the cathode and will cause the current flowing through V_1 to increase. This increase in current through V_1 will result in an increase in the voltage drop across R_2. Thus the voltage on the plate of the tube will become less positive with respect to ground; in other words, the plate voltage will swing in a negative direction.

When the signal applied by the generator reverses polarity, the grid will be made more negative and this will add to the bias voltage applied between the grid and the cathode of the tube. This will decrease the current flowing through the tube, which in turn will decrease the voltage drop across R_2. Therefore, the plate voltage on V_1 will swing in a positive direction.

Compare what we have in the plate circuit to the voltage applied to the grid by the generator. When the generator swings the grid positive, the plate voltage swings negative; when the generator swings the grid negative, the plate voltage swings positive. This means that the amplified signal voltage developed in the plate circuit of V_1 will be 180° out-of-phase with the input signal voltage eg.

The value of the coupling capacitor C_3 is chosen so that it has a low reactance within the range of signal frequencies we intend to amplify. As a matter of fact, the reactance of the capacitor is usually so low that for all practical purposes it acts like a direct connection for the signal voltages. Therefore, the voltage developed in the plate circuit of V_1 is applied directly to the grid of V_2 through the coupling capacitor C_3. The capacitor C_2 keeps the lower end of R_2 at signal ground potential. Often you will not find this capacitor in an amplifier; the actual capacitor is usually the output filter capacitor in the power supply. In any case, the signal voltage developed between the plate of V_1 and ground is coupled to the grid of V_2 through the coupling capacitor C_3 and to the cathode of V_2 through the cathode bypass capacitor C_4. Thus the amplified signal produced in the plate circuit of V_1 is coupled directly between the grid and cathode of V_2.

A TYPICAL TRANSISTOR CIRCUIT

Two R-C coupled transistor stages are shown in Fig. 2. Notice that both transistors are used in the common-emitter circuit and that both transistors are PNP transistors.

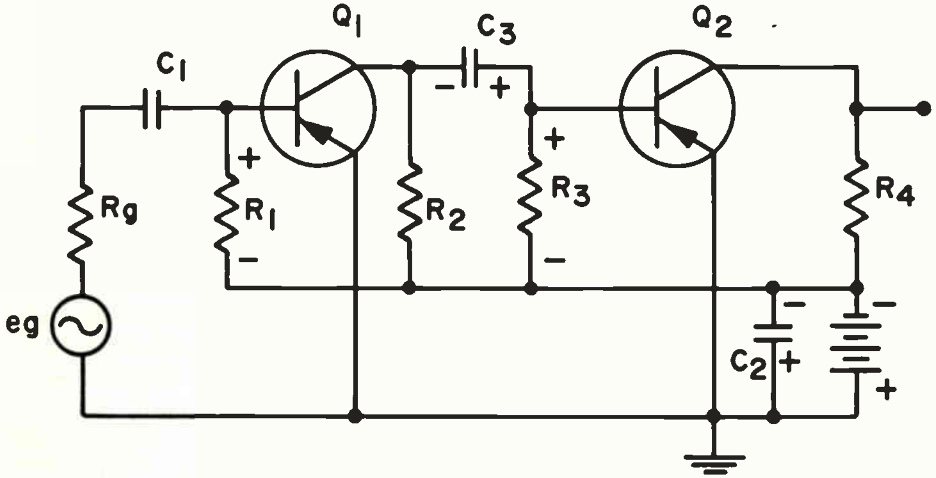


Fig. 2. Typical transistor R-C coupled stages.

The coupling components consist of R_2 (which is the collector load for Q_1), C_3 (the coupling capacitor) and R_3 (which is the base-bias resistor for Q_2).

Notice the capacitor across the battery. This is a large capacitor that effectively bypasses the battery as far as the signal is concerned, so that both the negative and positive sides of the battery are at ground potential in regard to the signal as well. In equipment designed for power line operation, the battery will be replaced by the power supply and C_2 will be the output filter capacitor.

Now let's review the operation of these two stages and see how a coupling network works. First, before any signal is applied, notice that the base of Q_1 and Q_2 connect to the negative side of the battery. The emitters of the two transistors connect to the positive side of the battery. This will place a forward bias across the emitter-base junction of the PNP transistor. However, the full battery voltage is not applied because there will be some small base current. A few of the holes crossing the emitter-base junction into the base will be filled by electrons that flow from the negative terminal of the battery through R_1 into the base of Q_1 and from the negative terminal of the battery through R_3 into the base of Q_2. Electrons flowing through resistors R_1 and R_3 will produce voltage drops across them having the polarity shown. The values of R_1 and R_3 are chosen so that the electrons flowing through them produce a voltage drop almost equal to the battery voltage. This will leave a forward bias across the emitter-base junction of only a few tenths of a volt.

Looking at the first stage, the emitter is connected to the positive side of the battery. This will pull electrons from the emitter, creating holes. The holes will be attracted by the negative potential on the base across the emitter-base junction, flow through the base and then across the base-collector junction and flow through the collector, where they will be filled by electrons coming from the negative terminal of the battery through R_2 to the collector. Thus we have a current flow through R_2 which will be governed by the number of holes reaching the collector of Q_1. This current flow through R_2 will result in a voltage drop across the resistor so that the negative potential on the collector of Q_1 will be somewhat less than the negative battery potential.

Holes and electrons move in the circuit for Q_2 in exactly the same way. Electrons pulled from the emitter of Q_2 create holes which flow through the transistor to the collector, where they are filled by electrons flowing through R_4.

When a signal voltage is applied by the signal source e_g the effective forward bias across the emitter-base junction of Q_1 is changed. Notice that we have shown a resistor in series with the generator. In the circuit for the vacuum tubes the generator will have some internal resistance, but since there is no current flow in the input circuit the resistance is of no consequence. However, in a transistor circuit there will be current flow in the input circuit and therefore the resistance affects the amount of signal actually reaching the transistor.

The capacitor C_1 is chosen to have a low reactance at signal frequency. Therefore the generator voltage is connected to the base of Q_1 through

R_g and C_1. The other side of the generator connects directly to the emitter.

When the end of the generator that connects to the base through the resistor and capacitor swings in a positive direction, the forward bias across the emitter-base junction of Q_1 will be reduced. This will cause the number of holes crossing the emitter-base junction and flowing to the collector to decrease. If fewer holes reach the collector then the number of electrons flowing through R_2 to fill these holes will go down. This means that the voltage drop across R_2 will decrease and the collector voltage will swing in a negative direction (in other words, closer to the negative battery potential).

When the generator polarity reverses so that the voltage applied to the base by the generator is negative, it will add to the forward emitter-base bias and cause the number of holes crossing the emitter-base junction to increase. This means that the number of holes reaching the collector will increase; therefore, the number of electrons flowing through R_2 to fill these holes must increase. The increase in current flow through R_2 will cause a greater voltage drop across this resistor. Thus, the potential at the collector of Q_1 will be less negative -- it will swing in a positive direction.

Notice the similarity between the transistor stage Q_1 and the vacuum tube stage V_1. In both cases, the signal is inverted; the amplified output signal voltage is 180° out-of-phase with the input signal voltage.

The amplified signal at the collector Q_1 is coupled to the base of Q_2, through the coupling capacitor C_3. This is similar to the arrangement used between the vacuum tubes in Fig. 1, but there is something quite different about the circuit. In the vacuum tube circuit shown in Fig. 1, there is no current flow through the grid resistor R_3 nor is there any current flowing between the cathode and grid of the tube. R_3 is a large resistor -- its purpose is to take care of any electrons that accidentally strike the grid of the tube and provide a path for these electrons back to ground. In most amplifier tubes this current flow through R_3 is so small it cannot be measured and we say for all practical purposes there is no current flow. This means that C_3 is coupling the signal into a very high impedance circuit.

In the circuit shown in Fig. 2, there is a current flow through R_3. However, the current flow through R_3 is small and the value of this resistor is comparatively large. Also, there is current flowing across the emitter-base junction. This is the movement of holes across this junction and since they flow across the junction with relatively little impedance with the forward bias on the transistor, we have, in effect, a low resistance circuit between the emitter and base. This is the circuit into which the capacitor C_3 must couple the signal. C_3 must be chosen so it has a low reactance compared to the input resistance of Q_2. Since the input resistance of Q_2 is quite low, the capacity of C_3 must be quite large in order to provide the low reactance coupling needed.

Often in transistor R-C coupled circuits, you will see electrolytic capacitors used as the coupling capacitor. In a circuit such as the one in Fig. 2 we have indicated the polarity with which an electrolytic capacitor should be connected.

The amplified signal from Q_1 is fed through C_3 to the base of the transistor. This causes the forward bias on Q_2 to vary, which in turn varies the number of holes flowing through the transistor and the electron current flow through R_4.

Q_2 will invert the signal 180° just like Q_1 did. Therefore, the signal voltage across R_4 will be in phase with the generator voltage. The signal voltage polarity will be inverted 180° by the amplifier stage Q_1 and will be inverted another 180° by the second amplifier Q_2 so that the output of the two-stage R-C coupled amplifier will be in phase with the input signal voltage. Of course, the same is true of the two-stage vacuum tube amplifier circuit.

When we discussed both the vacuum tube circuits shown in Fig. 1, and the transistor circuit shown in Fig. 2, we said that the coupling capacitor must have a low reactance. However, you know that the reactance of the capacitor varies with the frequency. Therefore, the reactance of these coupling capacitors must vary with the frequency; therefore, their effectiveness as coupling devices must also vary as the frequency varies. This fact has an effect on the limits of the frequencies which an amplifier of this type can amplify. There are other factors in the circuit which also limit the range of frequencies these amplifiers can handle and since these factors appreciably affect the performance of the amplifiers we will investigate them now.

FREQUENCY RESPONSE

When we talk about the frequency response of an amplifier we mean the ability of the amplifier to amplify signals of different frequencies. For example, if we say that the frequency response of an amplifier is flat from 100 cycles to 10,000 cycles we mean that signals from 100 cycles up to 10,000 cycles will receive the same amount of amplification. If we say that the frequency response of an amplifier falls off below 100 cycles and above 10,000 cycles we mean that signals having a frequency less than 100 cycles and signals having a frequency above 10,000 cycles do not receive as much amplification as signals between 100 cycles and 10,000 cycles receive.

Amplifiers can be designed to have a very wide frequency response -- in other words, they can amplify a wide range of frequencies. Often, however, it is not advantageous to have an amplifier that can amplify an extremely wide range of frequencies. For example, there is a limit to how low or to how high an audio frequency we can hear. Therefore, there is very little point in designing an amplifier that can amplify signals many times the frequency of the highest frequency we can hear. Certain economies can be realized by designing an amplifier that will amplify the desired or required frequency range and very little more. In addition to these economies, the design of the amplifier is usually simplified if we do not try to extend the frequency range it can amplify too much.

The factor that usually limits the frequency range of an amplifier most is the coupling network used between the various stages of the amplifier. In the circuits shown in Figs. 1 and 2 there are a number of factors that will limit the high-frequency and the low-frequency responses of these amplifiers. Let's look at these fac-

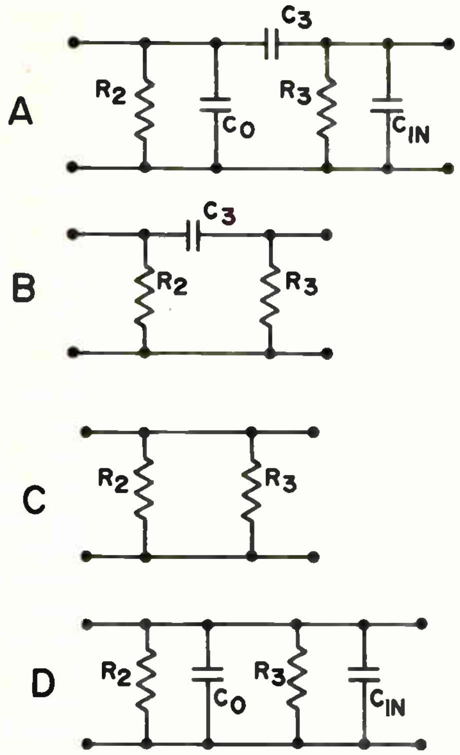


Fig. 3. Equivalent circuits of coupling network between V1 and V2 in Fig. 1.

tors and examine them in detail. We will consider the vacuum-tube circuit shown in Fig. 1 first.

In Fig. 3A we have shown a complete equivalent circuit of the coupling circuit used between V_1 and V_2. Notice that the plate resistor R_2 is shown in the circuit, as well as the coupling capacitor C_3 and the grid resistor R_3. In addition to these components we have also shown a capacitor marked C_O and a second capacitor marked C_{IN}. C_O represents the output capacity of V_1. This will be made up of the capacity in the tube itself plus wiring capacity in the circuit. C_{IN} is the input capacity of V_2. This will be the grid-to-cathode capacity plus any additional capacity added to the circuit by the wiring in the circuit. These capacitors are not shown on the diagram in Fig. 1, but they are present in the circuit and will affect the operation of the circuit.

At low frequencies, the output capacity of V_1 and the input capacity of V_2 are too small to appreciably affect the operation of the coupling network. We have shown the equivalent low-frequency circuit in Fig. 3B. Notice that the only components shown in this circuit are the plate resistor R_2, the coupling capacitor C_3, and the grid resistor R_3.

At some low frequency, the reactance of C_3 will be equal to the resistance of R_3. As you will remember, the capacitive reactance of C_3 increases as the frequency goes down. Even though the value of R_3 may be made quite large, if the frequency of the signal applied to the circuit is low enough a point will eventually be reached where the reactance of C_3 will be equal to the resistance of R_3. When this happens, C_3 and R_3 act as a voltage divider network so that only part of the voltage dropped across R_2 actually appears across R_3 and is fed to the grid and cathode of V_2.

By means of a vector diagram, we can see what happens to the voltage across R_3 when the reactance of C_3 is equal to the resistance of R_3. Remember that when a voltage is applied to a purely resistive circuit, a current will flow which will be in phase with the voltage. On the other hand, when a voltage is applied to a purely capacitive circuit, a current will flow that leads the voltage by 90°. When a voltage is applied to a circuit which has an equal resistance and an equal capacitive reactance, a current will flow in the

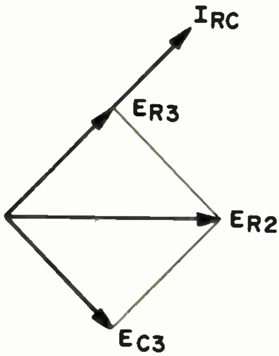


Fig. 4. Vector diagram of input and output voltages at low frequencies.

circuit that leads the voltage by 45°. From this information we can draw a vector diagram to show how the voltage across R_2 is divided across C_3 and R_3.

Fig. 4 is a vector representation of the voltage division. To draw this diagram the first thing we do is draw E_{R2}, which represents the voltage across R_2. Next, we draw the current I_{RC}, which represents the current that will flow through the network consisting of C_3 and R_3. We draw this current vector leading the voltage across R_2 by 45°. Now, we know that the voltage across C_3 will lag the current flowing through it by 90°. Therefore we can draw the vector E_{C3} lagging the vector I_{RC} by 90°. We also know that the voltage across R_3 will be in phase with the current flowing through it. This means that the voltage vector E_{R3} will fall on the vector I_{RC}. This will tell us the direction in which the vector representing the voltage across R_3 should point.

We know that the voltage across C_3 plus the voltage across R_3 must be equal to the voltage across R_2. Therefore, by drawing perpendiculars from the end of the vector representing the voltage across R_2 to the vectors representing the voltage across R_3 and C_3 we can determine the amplitude of the voltage across R_3 and across C_3. The perpendiculars are shown in Fig. 4. If we carefully measured the voltage vector E_{R3} we would find that it was equal to $.707 \times E_{R2}$. Similarly, the voltage across C_3 will be equal to .707 times the voltage across R_2. Thus, even though the capacitive reactance is equal to the resistance, the actual voltage that will appear across R_3 and be fed to the second stage will be slightly over 7/10ths the input voltage, where at first we might think that the voltage would be only half the input voltage.

The frequency at which the reactance of the coupling capacitor is equal to the resistance of the input resistor of the second stage is called the half-power point. The reason for this is that the current that will flow through the grid resistor will depend upon the voltage applied to it.

If the voltage at some frequency where the capacitive reactance of the capacitor is so small it can be ignored is 1 volt, then with the same amplitude signal across R_2, at the frequency where the reactance of the capacitor is equal to the resistance of the grid resistor, the voltage across the resistor will be .707 volts.

In this case, the current that will flow through the grid resistor will be .707 times what it would be when the voltage across the resistor was 1 volt. Therefore the power of the signal fed to the resistor will be equal to the voltage times the current which will be $.707E \times .707I = .5P$. Thus the power fed to the resistor will be one half the power at frequencies where the coupling capacitor can be ignored.

The equivalent diagram of the coupling circuit at mid-frequencies is shown in Fig. 3C. Here, the reactance of the coupling capacitor C_3 is so small compared to the resistance of R_3 that it can be ignored. Therefore for all practical purposes R_2 and R_3 are connected directly in parallel. R_3 is usually many times the resistance of R_2 so that it has very little effect as far as reducing the size of the plate load of V_1 is concerned. In the mid-frequency range the output capacity of V_1 and the input capacity of V_2 are too small to affect the operation of the circuit so they can be ignored. In this mid-frequency range, essentially all of the output signal developed by V_1 is fed to V_2 so that the coupling circuit operates with maximum efficiency.

At high frequencies the reactance of C_3 will be even smaller and can therefore be omitted from the circuit. However, the effects of the output capacity of V_1 and the input capacity of V_2 must be considered. They are shown in Fig. 3D, the equivalent circuit for high frequencies. Notice that once again R_2 and R_3 are, in effect, in parallel; also C_O and C_{IN} are likewise in parallel. Therefore the circuit acts as if there is one capacitor connected across the parallel combination of R_2 and R_3.

As the frequency of a signal increases, you know that the capacitive reactance decreases. At low and middle frequencies the capacitive reactance of C_O and C_{IN} in parallel is so high that it has no effect on the performance of the circuit. However, as the frequency increases, the capacitive reactance of this parallel combination goes down. At some high frequency, the capacitive reactance will eventually be equal to the resistance of R_2 and R_3 in parallel. At this point the effective plate load resistance on V_1 will be reduced and the output voltage developed by V_1 will also go down.

When the capacitive reactance of the parallel capacity is equal to the resistance of R_2 and R_3 in parallel, then half of the signal current developed by V_1 will flow through the resistance combination and the other half will flow through the capacitors. This means that the voltage developed in the output of V_1 will be reduced. By means of the vector diagram shown in Fig. 5, we can see exactly what happens.

First, we draw the current vector I which represents the signal current from V_1. Then we draw a voltage vector E_R to represent the voltage developed across the parallel combination of R_2 and R_3. We draw this vector in phase with the current because we know that the voltage developed across these resistors will be in phase with the current. Now we draw a voltage vector that lags the current by 90°. This vector represents the voltage developed across the capacitors. Since half the current is flowing through the resistance and half through the capacitance, the two voltages must be equal; therefore, we draw the voltage vector E_C equal to the voltage vector E_R and lagging by 90°. Now we perform the vector addition of these two voltages as shown in Fig.

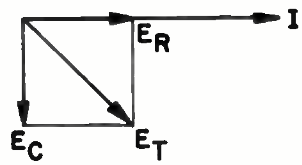


Fig. 5. Vector diagram of input and output voltages at high frequencies.

5. This gives us the total output voltage from V_1. Again, this voltage will be equal to .707 times the voltage developed in the mid-frequency range. This is the half-power point at the high-frequency end of the frequency range.

In addition to the amplitude of the signal voltage reaching V_2, dropping off both at high frequencies and at low frequencies, you should also notice that the phase of the voltage is changed. In the mid-frequency range, the voltage applied to V_2 will be in phase with the output voltage from V_1. However, at the low-frequency half-power point, the voltage fed to V_2 will lead the output voltage from V_1 by 45°. At the high-frequency half-power point, the voltage applied to V_2 will lag the output voltage from V_1 by 45°. In sound or audio amplifiers this phase shift is not particularly important, but in video amplifiers and television receivers this phase shift is important and can produce a smear in the picture. You will learn more about this later, as well as how this problem is overcome.

We have a somewhat similar situation in the transistor amplifier circuit shown in Fig. 2. However, here the situation becomes even more complicated because the transistors are low-impedance devices whereas tubes are high-impedance devices.

Fig. 6A illustrates the effective coupling circuit between Q_1 and Q_2 in Fig. 2. First, looking at Fig. 6A you see a resistor marked R_o. This resistance represents the output resistance of Q_1. In parallel with this we have shown R_2, the collector-load resistor. We have also shown C_O which represents the output capacitance of Q_1 plus any distributed capacitance that may be in the circuit. The coupling capacitor C_3 is shown as before; R_3 represents the resistance connected in the input circuit of Q_2. In parallel with R_3 we have shown another resistor marked R_{IN}. This resistance represents the input resistance or the base-emitter resistance of Q_2. In parallel with this combination we have capacitor C_{IN} which represents the input capacitance of Q_2. At some frequencies all of these parts affect the performance of the circuit.

Fig. 6B illustrates the equivalent low-frequency circuit of the coupling network. Here we have represented the output resistance of Q_1 and R_2 as a single parallel resistance. Since these two resistors are always in parallel we will use this representation in all of the equivalent circuits. Similarly, we have represented R_3

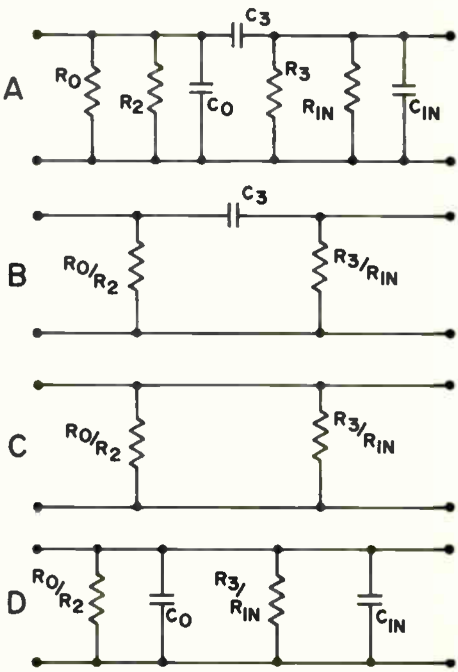


Fig. 6. Equivalent circuit of coupling network between Q1 and Q2 in Fig. 2.

and the input resistance of Q_2 as a single resistance. C_3 is the coupling capacitor and at low frequencies it will have a reactance that must be considered.

Since the input resistance of Q_2 is actually the parallel combination of R_3 and the input resistance to the transistor, we see that this resistance may be quite low. You already know that the input resistance of the transistor of a common emitter circuit is not particularly high; therefore, this resistance in parallel with R_3 will result in a comparatively low total resistance. Thus if we are to keep the half-power point at a reasonable frequency, we must use a large capacity for C_3. Capacitors many times those required for vacuum tube circuits are found in coupling networks between transistor stages. At some frequency the reactance of C_3 will be equal to the parallel resistance in the input circuit of Q_2 and then we will have the same voltage division we had in the vector circuit shown in Fig. 4. The amplitude of the voltage reaching the transistor Q_2 will be .707 of the amplitude reaching it in the mid-frequency range.

In Fig. 6C we have shown the effective mid-frequency range circuit. Here we simply have all of the resistances in parallel. Notice that since the input resistance of Q_2 is, in effect, in parallel with the output resistance of Q_1, the input resistance has an effect on the output resistance into which Q_1 is working. This means that in transistor R-C coupled amplifiers the input resistance of the second stage can actually affect the output that will be obtained from the first stage. This is not true in vacuum-tube coupled circuits because the input resistance of the second stage is so high it will not affect the output resistance of the first stage.

Fig. 6D shows the equivalent high-frequency circuit. We have shown the two resistances and capacitances separately, but actually they could be lumped together so that we have, in effect, a single resistance and a single capacitance. As in the case of the vacuum-tube circuit, at some frequency the capacitive reactance of the parallel capacity will be equal to the resistance of the parallel resistors. When this happens, the output voltage developed by Q_1 will fall to .707 of the voltage developed in the mid-frequency range. The vector representation will be the same as that given in Fig. 5.

We mentioned that the coupling capacitor in a transistor circuit must be larger than in a tube circuit in order to keep the low-frequency half-power point at a reasonably low frequency. In other words, we have an additional problem in the transistor circuit because of the low input resistance of the second transsistor. At the high frequency end we have a somewhat different situation. Since the transistors are already low-resistance devices, the value of the output and input capacitances can be somewhat larger than in a vacuum-tube circuit before they cause appreciable difficulty. If they happen to be the same as those in a vacuum tube circuit, then the high-frequency half-power point in a transistor circuit will be somewhat higher than the vacuum-tube circuit because the vacuum-tube circuits are high-resistance circuits and are quickly loaded by a capacitive reactance. The transistor circuits, on the other hand, are low-resistance circuits and the high-frequency half-power point will not be reached until a somewhat higher frequency.

CASCADE AMPLIFIERS

At the half-power point we say that the gain of an amplifier is .707 times the gain of the amplifier middle frequency. We also express this quite frequently as the percentage and say that the gain is 70.7% of the gain in the middle frequency range. This drop-off in gain due to the coupling network between a single stage is not too troublesome when only one network is concerned, but sometimes we have a number of amplifiers called cascade amplifiers and here the problem becomes much greater.

Cascade amplifiers are just a number of amplifiers connected together. A diagram of four cascade amplifiers is shown in Fig. 7. Here the input signal is fed to amplifier 1, amplified by it and then fed to amplifier 2, amplified further and fed to amplifier 3, where it is amplified still more, fed to amplifier 4, and amplified again.

We have not shown actual circuits in this diagram but we have presented each stage as a block to simplify the diagram. Each of these amplifiers could be either a vacuum tube amplifier or a transistor amplifier.

If the gain of each amplifier is 10 in the middle frequency range of about 100 cycles to 10,000 cycles, then the over-all gain of the first two stages is 10 × 10, or 100. The gain of the first three stages is 100 × 10 or 1000, and the gain of all four stages is 10,000. Thus, in the middle frequencies where the gain of each stage is 10, the total gain of this system is 10,000.

Now let's assume we are interested in amplifying signals as low as 75 cycles and at that frequency, the gain of the amplifier drops to 70.7% of what it is in the middle frequency range. Here the gain of each stage would be 7.07. Now the gain of the first two stages is 7.07 × 7.07 = 50. Now notice that the voltage gain of the two stages is only half of what it was at the middle frequencies.

The gain of the first three stages will be 50 × 7.07 or a little over 350, and the gain of all four stages will be 7.07 × 7.07 × 7.07 × 7.07 = 2500. Notice that whereas the gain of each stage has fallen to 70.7% of the gain at the middle frequencies, the over-all gain of the amplifier is only 25% of what it was at the middle frequencies! Thus, you can see that although the decrease in gain to 70.7% is not too big a problem in a single coupling network, if we have a number of stages used together, this fall-off in output is cumulative, so that in the four stages coupled as in Fig. 7, we have a gain of only 25% of the gain we had at the middle frequency. Of course, this much drop in response could not be tolerated and we would

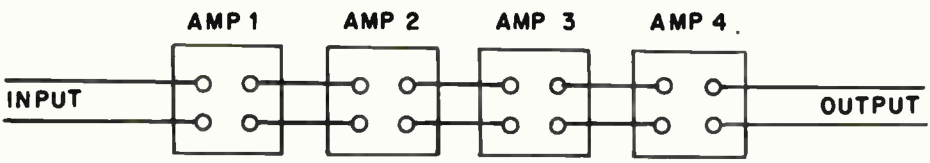


Fig. 7. Block diagram of cascade voltage amplifiers.

have to design each stage a little better so that within the range of frequencies we wanted to amplify, the gain in each stage would remain almost constant. In the example we have given, this problem can be overcome by using larger coupling capacitors.

SUMMARY

Resistance-capacitance coupled voltage amplifiers are the most important type of audio amplifier you will encounter. You should be able to draw from memory the circuits shown in Fig. 1 and in Fig. 2. You will use these circuits over and over again and the sooner you know exactly how the circuits are connected, the faster you'll reach a point where you'll understand their operation completely.

In the two circuits the first stage amplifies the input signal voltage. The amplified signal voltage is developed across the load resistor R_2. This amplified voltage is then fed through the coupling capacitor C_3 to the following stage. As long as the reactance of C_3 is low compared to the input resistance of the following stage all of the signal developed across R_2 reaches the input circuit of the following stage. However, at low frequencies the reactance of C_3 becomes appreciable compared to the input resistance of the following stage and then part of the signal is lost across C_3 so that the entire signal does not reach the input of the following stage. Thus in a low-frequency range the gain of the two-stage amplifier begins to fall off.

At some high frequency the output capacitance of the first stage and the input capacitance of the second stage begin to have a low enough reactance to effectively shunt the circuit and reduce the signal voltage developed across the load resistor R_2. When this happens, the gain of the amplifier begins to fall off as it does at low frequencies due to the reactance of the coupling capacitor. Because the input resistance of a transistor amplifier is much lower than that of a tube amplifier, a much larger coupling capacitor is required in the low-frequency region in the transistor circuit than is required in a tube circuit.

In the case of the transistor amplifier there was a tendency for the shunting capacity to have a less serious effect on the gain of the amplifier than it had in the case of the tube amplifier. This was due to the fact that the output resistance of the first transistor and the input resistance of the second stage were already low; therefore, a much larger capacitance was required to have the same loading effect. The input resistance of the second stage in a transistor amplifier had an appreciable effect on the gain of the first stage -- this is not true in the case of a vacuum tube amplifier; the input resistance of the second stage in the latter type is so high that it will not have any effect on the value of the load resistance in the plate circuit of the first stage.

Be sure that you understand resistance-capacitance amplifiers thoroughly before leaving this important section of the lesson. After you're sure that you understand the amplifiers, check yourself by doing the following self-test questions.

SELF-TEST QUESTIONS

(a) What is the purpose of R_1 in the circuit shown in Fig. 1?

(b) What is the purpose of C_1 in Fig. 1?

(c) With respect to the input signal voltage, what will the polarity of the amplified signal voltage across R_2 be?

(d) Is the voltage drop across R_1 in Fig. 2 used as the forward bias on Q_1?

(e) In the circuit shown in Fig. 2, do the holes reaching the collector flow through the collector load resistor R_2?

(f) When the input signal from the generator in Fig. 2 has a polarity which causes the end that connects the emitter to be negative and the end that connects to and through C_1 to the base to be positive, what happens in the transistor?

(g) Why does the input resistance of Q_2 in Fig. 2 have an effect on the gain of Q_1?

(h) In the circuits shown in Figs. 1 and 2, what part primarily limits the lower frequency limit of the two amplifiers?

(i) In the same circuits, what limits the high frequency limit to which the amplifiers can be used?

Transformer-Coupled Voltage Amplifiers

In the early days of radio, the vacuum tubes manufactured were relatively crude, and you couldn't get a great deal of gain from them. Voltage amplifiers with vacuum tubes frequently used transformer coupling between the tubes because the gain of the stage could be increased considerably by using a step-up transformer. For example, suppose the tube had a gain of five and you used a step-up transformer between the tube and the following stage, and suppose that the transformer had a turns-ratio of 1 to 3; this meant that the transformer would step-up the signal voltage fed to the primary by a factor of 3. Therefore, the total gain obtained with the tube and transformer would be equal to 5 × 3, or 15.

With modern vacuum tubes it is comparatively easy to obtain a gain of nearly 100 in R-C coupled circuits. Since R-C coupling is much more economical than transformer coupling, you will not find transformer coupling between voltage amplifier stages in modern equipment. Therefore, we'll take only a quick look at a vacuum-tube voltage amplifier using transformer coupling so you will know what it looks like, in case you run into it in any older equipment you might service, and also because transformer coupling is still used in some other applications.

In transistor circuits an entirely different situation exists. We have already pointed out that the input resistance of the second stage in a two-stage transistor amplifier has an effect on the output resistance of the first stage. We pointed out that the low input resistance of the second stage will, in effect, reduce the load resistance of the first stage, which in turn will reduce the amplitude of the output signal from the first stage. This problem can be overcome by using transformer coupling between the two stages. The transformer serves as an impedance-matching device and prevents the low input resistance of the second stage from reducing the load resistance of the first stage. Today, transformer coupling is much more important in transistorized circuitry than in vacuum circuity, so we'll spend more time discussing this type in transistorized circuits. However, let's go ahead and look at a transformer-coupled vacuum tube amplifier first.

VACUUM-TUBE CIRCUITS

Fig. 8 illustrates a typical two-stage vacuum-tube amplifier using transformer coupling between the stages. The operation of the circuit is comparatively simple. Resistor R_1 provides bias for V_1, and C_1 is the cathode bypass which provides a low-impedance bypass around the resistor for the signal. Resistor R_2 similarly provides bias for the second stage, V_2, and it is bypassed for the signal by capacitor C_2.

The input signal voltage causes the

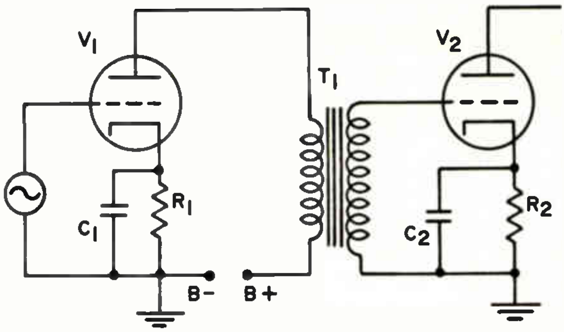


Fig. 8. Two transformer-coupled vacuum tubes.

potential between the grid and cathode to vary. This causes the plate current flowing through V_1 to vary. This varying current flows through the primary of T_1 and sets up a varying magnetic field which induces a voltage in the secondary of T_1. T_1 is a step-up transformer so that the voltage across the secondary will be greater than the voltage across the primary.

The amplified and stepped-up voltage appearing across the secondary of T_1 is applied between the grid and cathode of V_2 where it will receive further amplification.

At first glance it might appear that the transformer-coupled two-stage amplifier has many advantages over the R-C coupled amplifier. However, as we pointed out before, the step-up feature of the transformer is not needed with high-gain vacuum tubes available to use in modern equipment. In addition, a transformer is much more costly than the two resistors and capacitor needed for an R-C coupled network. Furthermore, most transformers are quite frequency-sensitive and therefore the signal voltage reaching V_2 will vary appreciably with frequency. Some of the problems of frequency discrimination in the transformer can be overcome by careful design, but this in turn increases the cost of the transformer. Because of the disadvantages of the transformer and the availability of high-gain tubes, this circuit is not used in modern equipment.

TRANSISTOR CIRCUITS

A transistor voltage amplifier using transformer coupling is shown in Fig. 9. Notice that both transistors are connected in common-emitter circuits.

From a standpoint of getting maximum gain from a transistor amplifier, transformer coupling is actually the best arrangement that can be used. You will remember that the input impedance of a transistor is quite low. We also pointed out that when two transistors are coupled together by means of resistance-capacity coupling, the low impedance of the second transistor actually loads the first transistor. The two transistors are not matched. By means of a transformer the first transistor can be matched to the second transistor so that the undesirable effect of having the input circuit of the second transistor load the output circuit of the first transistor can be avoided. The transformer serves as an impedance-matching device to match the comparatively high impedance of the first transistor to the low input impedance of the second transistor.

Looking at Fig. 9 we see that we have a transformer, T_1, in the input. This transformer is a step-down transformer and matches the low input impedance of the first transistor to the preceding circuit. C_1 is a blocking capacitor. It is needed to keep the transformer secondary from shorting out the forward bias placed across the emitter-base junction. The bias on the junction is obtained from the battery by means of a voltage divider network consisting of both R_1 and R_3. This network will bias the base negative with respect to the emitter. Notice that the 3K resistor R_1 is connected directly across the input circuit. As for the signal, it is connected directly from the base to the emitter. The 4 mfd capacitor C_3 provides an effective signal bypass between the emitter and ground.

The resistor R_2 is placed in the circuit to prevent thermal runaway. You will remember we mentioned previously that minority carriers crossing the collector-base junction may increase the forward bias across the emitter-base junction. This causes the current flow through the transistor to increase, heating the collector-base junction and its resistance, thereby causing still more minority carriers to cross the junction and increase the forward bias still further. R_2 prevents this from happening because if the emitter current increases, the voltage across R_2 increases, and this voltage subtracts from the forward bias across the emitter-base junction. This tends to keep the current through the transistor constant.

The signal is applied through the 4 mfd capacitor to the base of the first transistor and from the other end of the transformer through C_3 to the emitter. This causes the number of holes crossing the emitter-base junction to vary which in turn causes the number of holes reaching the collector to vary. Then the negative flow from the negative ter-

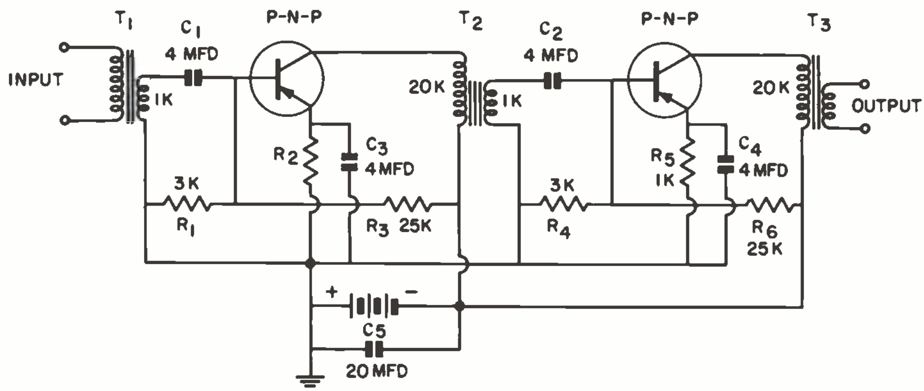


Fig. 9. Transformer-coupled voltage amplifier using transistors.

minal of the battery through the primary of T_2 will vary. The varying current through the primary of the transformer produces a magnetic field which induces a voltage in the secondary of the transformer. This in turn produces a varying voltage between the base and emitter of the second transistor.

Notice that the primary of T_2 is marked 20K and the secondary is marked 1K. The output impedance of a transistor in a common-emitter circuit is about 20K. The input impedance is about 1K. By means of a step-down transformer these impedances can be matched. You will remember that maximum power transfer from a generator to the load occurs when the load impedance is matched to the generator impedance. The first transistor acts more or less as a generator and maximum power is transferred from it to the second transistor when the input impedance of the second transistor is matched to the output impedance of the first transistor.

You might wonder why we are concerned about matching impedances to get maximum power transfer when the circuit shown in Fig. 9 is supposed to be a voltage amplifier. Remember, however, that transistors are current-operated devices. Therefore by matching impedances we can get maximum current variation in the emitter-base circuit of the second transistor. This in turn will result in maximum current variation in the collector circuit. The higher the current variation through the load of the second transistor the greater the voltage amplification obtained will be.

A two-stage transformer-coupled amplifier such as is shown in Fig. 9 will yield far more voltage gain than the two-stage R-C coupled amplifier shown in Fig. 2. This might at first lead you to believe that transformer coupling should be used in all transistor voltage amplifier circuits. However, this is not true because the transformers needed for the transformer coupling are much more expensive than the two resistors and capacitor needed in R-C coupled circuits. As a matter of fact, since transistors themselves are relatively inexpensive, it is usually more economical to use a three-stage R-C coupled voltage amplifier than it is to use a two-stage transformer-coupled voltage amplifier.

A three-stage R-C coupled amplifier is capable of giving about the same gain as a two-stage transformer-coupled amplifier. In addition, it is more difficult to get good frequency response using transformer coupling than it is using R-C coupling. Therefore, in applications in which we are concerned about the frequency response and in which we are trying to keep the cost down it is usually advantageous to use a three-stage R-C coupled amplifier. However, in some applications you will find transformer-coupled voltage amplifier circuits.

If you have to replace a transformer in a transformer-coupled voltage amplifier, it's important to use a replacement transformer having the same turns-ratio as the original transformer. The transformer is primarily an impedance-matching device and if you do not use a transformer with the same turns-ratio as the original, the output impedance of the first transistor will not properly match the input impedance of the second transistor.

SUMMARY

Since it is uneconomical, transformer coupling between voltage-amplifier vacuum-tube stages is no longer used. You will find it only in very old equipment. The transformer used between vacuum-tube voltage-amplifier stages is usually a step-up transformer. Transformer coupling is used in some transistor amplifiers. In this application, the transformer is a step-down transformer and it is primarily an impedance-matching device. By using a transformer in transistor voltage amplifiers, the high output impedance of the first transistor can be matched to the low input impedance of the second transistor. Transistor voltage amplifiers are capable of higher gain than R-C coupled amplifiers, but by using an extra stage the same gain can be obtained from R-C coupled amplifiers.

SELF-TEST QUESTIONS

(j) Why was transformer coupling used between vacuum tubes in the early days of radio?

(k) Why are transformer-coupled voltage-amplifier stages using tubes no longer used?

(l) Why is it better to use transformer coupling in transistor-voltage amplifier stages?

(m) Why is transformer coupling seldom used between transistor-voltage amplifier stages?

(n) What must you watch if you have to replace a transformer in a transformer-coupled transistor-voltage amplifier?

Single-Ended Power Amplifiers

The primary purpose of most low-frequency amplifiers is to build up an audio signal to a sufficient amplitude to drive the loudspeaker. A loudspeaker must be driven by a power amplifier. The voltage amplifiers we have discussed previously are used to amplify a low-voltage audio signal to a sufficient amplitude to drive a power amplifier, which in turn will drive the loudspeaker.

There are a number of different types of power amplifiers that can be used to drive a loudspeaker: the single-ended power amplifier, which is always a Class A power amplifier, and the double-ended power amplifier, which may be a Class A, Class AB or Class B power amplifier. In this section of the lesson we are going to discuss single-ended power amplifiers. We will discuss both vacuum-tube power amplifiers and transistor power amplifiers.

Today, there are millions of radio and television receivers in use with low-frequency single-ended power amplifiers using vacuum tubes. You can be sure that you will run into this type of amplifier to service. In addition, there are many single-ended power amplifiers using transistors. This type of power amplifier is found in modern automobile receivers as well as in many portable receivers, in some table model radios and in some transistorized television receivers. It is quite likely that in the future there will be fewer vacuum tube stages around to service, but at the present there are probably many more power amplifiers using vacuum tubes than transistors. We will go into the vacuum-tube circuit first, and then into the transistor power amplifiers later.

VACUUM-TUBE CIRCUITS

A typical single-ended Class A power amplifier using a beam power tube is shown in Fig. 10. Practically all modern low-frequency power amplifiers use beam power tubes. This stage is driven by the amplified signal voltage from the last voltage amplifier applied to R_1 through C_1. It produces audio power to drive the loudspeaker. Notice that in general this circuit is not too different from that of a voltage amplifier.

The tubes made for use in power amplifier stages are designed especially for this type of service. These tubes normally draw a much higher plate current than the tubes in voltage amplifiers. A typical tube used in a voltage amplifier may draw a plate current somewhere between about 1 and 10 milliamperes, whereas tubes designed for use in power amplifiers will usually draw a plate current of 50 milliamperes or more.

In the circuit shown in Fig. 10, R-C coupling is used between the preceding voltage amplifier and the power amplifier. This type of coupling can be used between a Class A power amplifier and the last voltage amplifier because no grid current flows. The power amplifier shown in Fig. 10 is designed to operate as a

Class A amplifier by the voltage developed across the cathode-bias resistor R_2. You will remember that in a Class A amplifier, sufficient bias is applied to the tube so that the grid voltage is halfway between zero and cut-off voltage. In operation, the grid should not be driven positive, and it should not be driven into the region where plate current is cut off. As long as the grid is not driven positive, there will not be a grid current flow.

Notice that R_2 is bypassed by a capacitor, C_2. This capacitor is usually an electrolytic capacitor -- in many receivers you will find that this capacitor is in a common can along with the filter capacitors used in the receiver. It is usually a comparatively low-voltage capacitor because the voltage across it (which is equal to the voltage across R_2) is not too high. An electrolytic capacitor is used because the resistance of R_2 is usually only a few hundred ohms. A large capacitor is required to bypass this resistor -- otherwise, the reactance of the capacitor becomes so high at low frequencies that the resistor is not bypassed effectively. The purpose of this capacitor is to hold the cathode at signal ground potential, as we mentioned previously, so that the input signal is, in effect, applied directly between the grid and cathode of the tube.

You will remember that to get maximum power transfer from a generator to a load, the load impedance must be matched to the generator impedance. In this circuit, the tube acts as the generator. However, a beam power tube has a very high plate impedance. Therefore, it is not practical to try to match the plate impedance of the tube to the speaker. Instead, manufacturers of the tubes usually specify a load impedance into which a tube should work for best results. The output transformer is designed so that the tube will, in effect, see this load impedance. This is accomplished by means of a step-down transformer, T_1, which matches the load impedance of the tube to the low impedance of the speaker voice coil. Another way of looking at this transformer is from the speaker end of the circuit. Most speakers have a comparatively low impedance; the speakers in most radio and television receivers have a voice coil

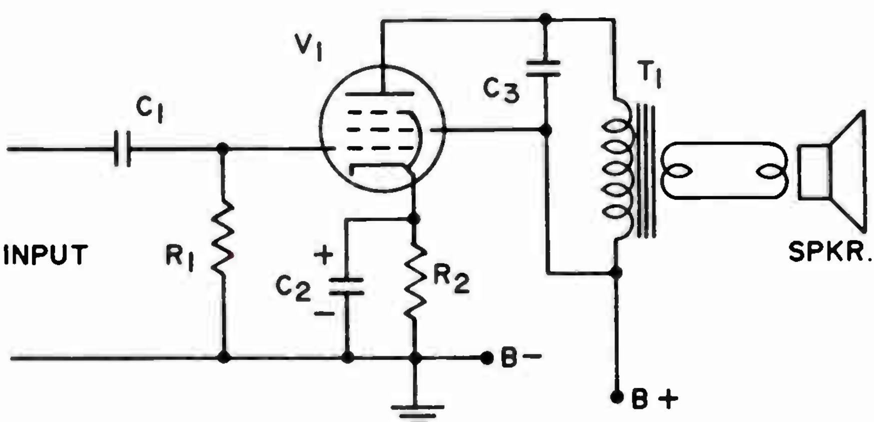


Fig. 10. Schematic diagram of single-ended Class A amplifier using a beam power tube.

impedance of 3.2 ohms at a frequency of 400 cycles. The output transformer is designed to step up this impedance so that the tube sees the load it works into best. Most beam power tubes work into a load of somewhere between 2,500 ohms and 10,000 ohms. Therefore, if we look at the circuit from the speaker end, the transformer is a step-up transformer because it steps up the low impedance of the speaker to the higher impedance of the tube. If we look at the circuit from the tube end, on the other hand, the transformer is a step-down transformer because it steps the high impedance of the tube down to the low impedance of the speaker. The important thing for you to realize is that the primary winding of the transformer, which connects to the plate circuit of the tube, has far more turns on it than the secondary winding that is connected to the voice coil of the speaker. Since the primary winding has more turns than the secondary, we usually refer to this as a step-down transformer.

Capacitor C_3 is called a plate bypass capacitor. This capacitor is used to prevent oscillation. Actually, the capacitor provides a low-impedance path for high-frequency signals and it reduces the possibility of a high-frequency oscillation in the stage. These high-frequency signals in the plate circuit of the tube are shunted to ground through the capacitor while low frequency signals must flow through the primary of T_1. You might wonder how the high-frequency signals are shunted to ground through C_3, since C_3 connects between the plate and screen of the tube. The screen of the tube connects directly to B+. The output of the power supply will have a large electrolytic capacitor as the output filter capacitor. This capacitor provides a low impedance path from B+ to ground as far as the audio signal is concerned. Therefore, B+ is effectively at ground as far as the signal is concerned. Hence the bypass capacitor C_3 (which connects from the plate to the screen of the tube) effectively bypasses the high-frequency signal to ground.

In some receivers you will find C_3 connected directly between the plate of the tube and ground. The disadvantage of this arrangement is that C_3 must then be capable of withstanding the high dc plate voltage applied to the tube. By using the circuit shown in Fig. 10 and connecting C_3 between the plate and screen of the tube, the only dc voltage the capacitor must be able to withstand is the voltage drop across the primary winding of T_1. This voltage is usually nominal (somewhat less than 10 volts) so that C_3 does not have a high dc potential placed across it.

Sometimes you will see a triode tube used as the power output tube. A triode tube has the advantage over a pentode or a beam power tube, in that it usually produces less distortion. However, the disadvantage of the triode tube is that it requires a very high driving power. The power gain in a triode tube is relatively low compared to that of a beam power or a pentode tube. As a result, triode tubes have in general disappeared as power amplifiers in favor of beam power and pentode tubes.

TRANSISTOR CIRCUITS

A schematic diagram of a single-ended Class A amplifier using a PNP

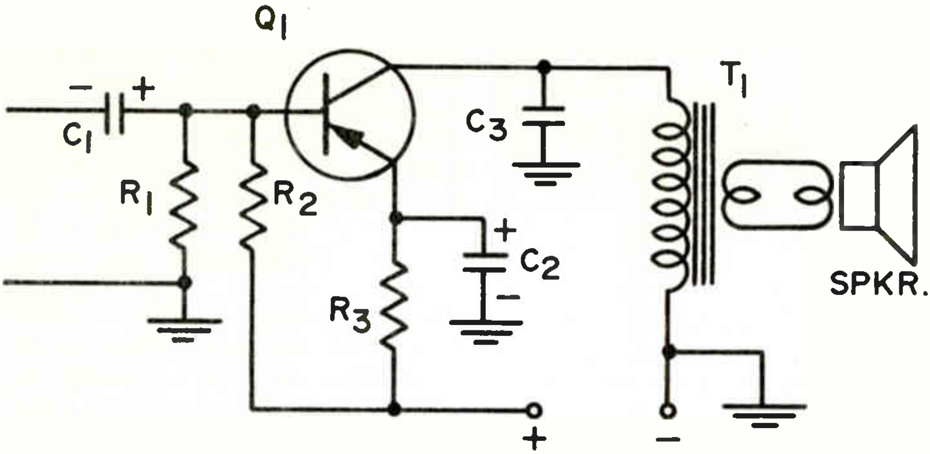


Fig. 11. Schematic diagram of single-ended Class A amplifier using a PNP transistor.

transistor is shown in Fig. 11. Notice that, in many respects, this circuit is similar to the vacuum tube circuit shown in Fig. 10.

The circuit we have shown in Fig. 11 is a common-emitter circuit and the output transformer is connected in the collector circuit. Again, the output transformer is an impedance-matching device. Its purpose is to match the speaker impedance to the output impedance of the transistor. Notice that capacitor C_2 is an electrolytic capacitor. The value of R_3 is comparatively low, so in order to hold the emitter at signal ground potential for all signal frequencies, the capacitor must be large. As in the case of a vacuum-tube circuit there will be some low frequency where the reactance of this capacitor becomes appreciable and reduces the gain of the amplifier at that frequency.

The input coupling capacitor C_1 (in the circuit shown) is an electrolytic capacitor. Again, remember that an electrolytic is required in the input of the stage because the input resistance of the transistor is very low and unless a large capacitor is used, there will be considerable loss in gain at low frequencies.

Once again, C_3 is a bypass capacitor. Its purpose is the same as that of the plate bypass capacitor in Fig. 10. It is used to prevent oscillation in the collector circuit and to provide a low-impedance path from the collector to ground for high-frequency signals in the collector circuit.

In a Class A stage of this type, the emitter-base junction is biased so that the transistor operates approximately in the midpoint of its characteristic curve. The bias is adjusted so that the signal does not overcome the emitter-base forward bias at any time and the current flow from the emitter to the collector is never cut off.

Transistor power amplifiers of this type are quite widely used in automobile receivers. In most automobiles the negative side of the battery is grounded. This means that the collector of the transistor is almost at ground potential. It is at ground potential less the voltage

Fig. 12. A power transistor.

drop across the primary of T_1. As a result, very little insulation is required between the collector and the receiver chassis. Power transistors resemble the transistor shown in Fig. 12. They are usually bolted to the chassis to provide a heat sink, or a means of getting rid of the heat dissipated by the transistor. Very little insulation is required between the transistor case and the chassis in a circuit of this type because there is very little voltage between the chassis and the collector. In most power transistors, the collector connects directly to the case of the transistor; therefore, the circuit provides a simple way of getting rid of the heat developed in the transistor and at the same time keeps the insulation requirements between the transistor case or collector to a minimum.

Of course, in a circuit of this type NPN transistors can be used as well as PNP transistors. For that matter, in automobile receivers where the battery voltage is only 12 volts, it is comparatively simple to insulate NPN transistor collectors from the chassis so that there will be no voltage breakdown, and at the same time provide adequate heat dissipation from the transistor.

Single-ended transistor amplifiers are also used in radio and television receivers that are operated from the power line. However, in portable-type equipment that operates from a dry cell-type battery, it is more economical to use push-pull transistors in Class B circuits. You will see this later.

SUMMARY

Class A power amplifiers are important because they are used in all types of equipment, especially in radio and television receivers. The power amplifier is a stage that converts the amplified signal to a signal with sufficient power to drive a loudspeaker. It consists of a tube or transistor, the output transformer, and the loudspeaker. The loudspeaker is the load; the output transformer is designed to match the speaker impedance to the output impedance of the tube or transistor.

You should be sure that you understand what a Class A power amplifier is and what it is supposed to do. If you go into radio and TV service work, you will run into an amplifier of this type in almost every radio or television set you will repair. If you plan to be a technician in a broadcast station you will find Class A power amplifiers in both radio and television equipment. If you intend to go into industrial electronics, you may be in a plant where a public address system is used. The chances are that you will find Class A amplifiers in this equipment. Regardless of what type of electronics work you do, you will encounter this type of amplifier.

SELF-TEST QUESTIONS

(o) What is the purpose of a power amplifier in a radio receiver?

(p) What type of tube can you expect to find in the power amplifier stage of most modern radio and television receivers?

(q) What is the purpose of the bypass capacitor used in the plate circuit of power amplifier tube circuits?

(r) In a circuit such as the one shown in Fig. 10, how can low-frequency distortion due to the reactance of the cathode bypass capacitor be kept to a minimum?

(s) What is the advantage of using a PNP transistor power amplifier (such as in Fig. 11) in an automobile receiver?

(t) Are you likely to find a single-ended or a push-pull power output stage in portable receivers which operate from batteries?

Push-Pull Power Amplifiers

In many applications it is impossible or at least uneconomical to try to obtain the required power from a single-ended stage. In such cases a double-ended power amplifier called a push-pull amplifier is used. A push-pull amplifier is particularly useful because it generates no second-harmonic distortion. Because of the circuit arrangement, any second-harmonic distortion produced is cancelled within the stage itself. Second-harmonic distortion is a signal having a frequency equal to twice the frequency of the signal to be amplified. Due to the curve in the characteristic curve of both tubes and transistors, single-ended stages produce a certain amount of second-harmonic distortion. However, in push-pull amplifiers, any second-harmonic distortion that is produced will be cancelled in the stage itself.

Push-pull stages may use either tubes or transistors, and you will study both types. Also, push-pull stages may be operated as Class A amplifiers or as Class B amplifiers. Often, you will find push-pull pentode or beam power stages operated as Class AB amplifiers. We will study and review some of the conditions of these various types of operations.

CLASS A AMPLIFIERS

A typical Class A push-pull amplifier using beam power tubes is shown in Fig. 13. Transformer T_1 can be a step-up transformer to provide additional drive between the grid and cathode of each tube. Notice that the cathode bias resistor R_1 is not bypassed. A bypass capacitor is not necessary because the current flowing through this resistor remains essentially constant. When the input signal drives the grid of V_1 in a positive direction so that the current flow through this tube increases, the signal will at the same time drive the grid of V_2 in a negative direction so that the current through that tube decreases. If the tubes are operated on the linear portion of their characteristic curves, the increase in plate current in V_1 should be offset by the decrease in plate current in V_2.

C_1 and C_2 serve as plate bypass capacitors for the two tubes. These capacitors prevent high-frequency oscillations which might cause the tubes to draw excessive current and burn out the primary of the output transformer.

Notice that the B+ for the screen of the tubes is obtained from the output filter capacitor. The B+ at this point will receive maximum filtering so that it is essentially pure dc. At the same time, the center tap of the output transformer, T_2, through which plate voltage is supplied to the two tubes, connects to the input filter capacitor. The dc voltage at this point will be somewhat higher than the voltage at the output filter capacitor due to any voltage drop that might occur in the filter choke or filter resistor used in the power supply. In addition, there will be considerably more hum voltage present at this point. At first, you might think that this would cause hum in the output. However, there are two factors which prevent hum in a circuit of this type. First, any hum current would flow through the two halves of the output transformer

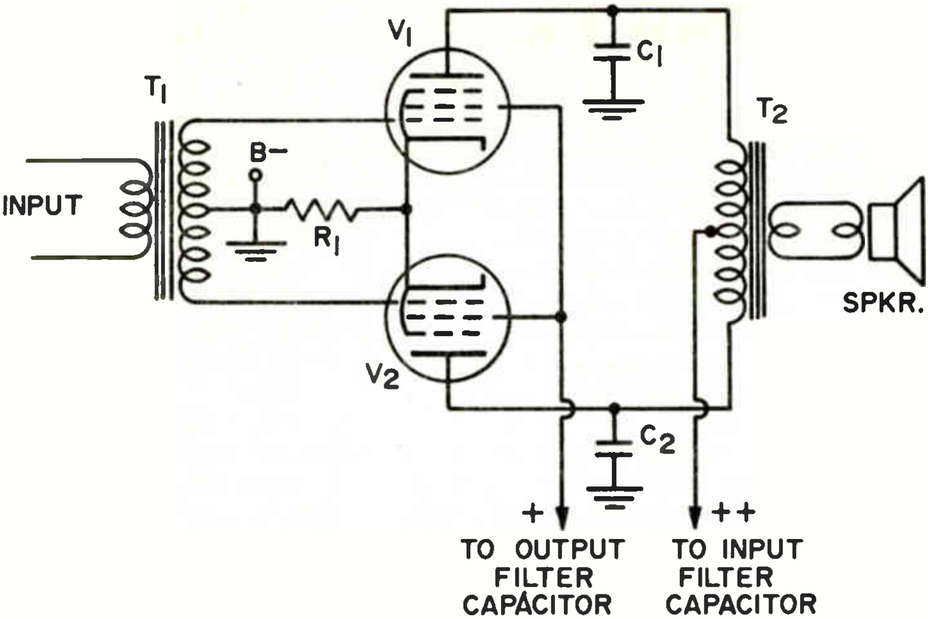


Fig. 13. Class A push-pull beam power tubes.

T_2 in opposite directions and hence tend to cancel. In addition, if there is some hum voltage on the plates of the two output tubes, remember that in pentodes and in beam power tubes the plate current depends very little on the plate voltage. It is determined primarily by the grid and screen voltages. Therefore, if there is some variation in the plate voltage on the tubes due to hum, this will not cause any appreciable change in plate current. Hence, the current flowing through the primary of T_2 remains essentially constant, and no hum will be fed to the speaker.

A push-pull Class A power amplifier using NPN transistors is shown in Fig. 14. Notice that, in many ways,

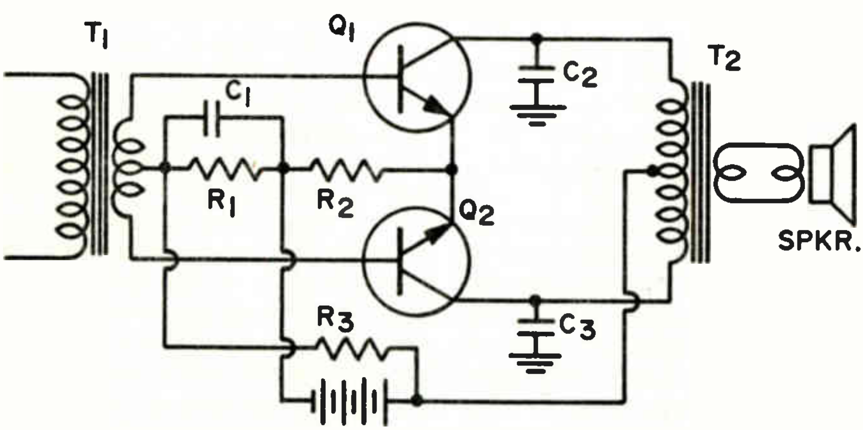


Fig. 14. Class A push-pull NPN transistors.

the circuit is similar to the vacuum-tube circuit in Fig. 13. The transformer T_1 is a step-down transformer; you will remember that the input resistance (or impedance, as it's more correctly called when referring to a signal) of transistors in a common-emitter circuit is quite low. Thus, T_1 serves as a matching device. Forward bias for the emitter-base junction is obtained by means of a voltage-divider network made up of R_1 and R_3 connected across the battery.

Resistor R_2, connected between the negative side of the battery and the emitter, is for bias stabilization purposes. In the event that the current through the transistor starts to increase due to a temperature rise in either transistor, the current through R_2 will increase and reduce the bias to hold the current through the transistors essentially constant.

Capacitors C_2 and C_3 are bypass capacitors that are used to prevent high-frequency oscillation in the output circuit. In some circuits the capacitors will be connected as shown, while in others a single capacitor connected directly between the two collectors may be used.

For simplicity, we have represented the power supply in the circuit as a battery. Of course, in a portable receiver or in an automobile receiver the actual power source will be a battery. However, in equipment designed for use in the home (where it will be operated from a power line) a power supply will replace the battery. The power supply will use a rectifier and a filter circuit similar to those you have already studied. It might also have a step-down transformer if the transistors are of the low-voltage type. However, many transistors are designed for high-voltage operation and equipment using this type of transistor will not require a transformer in the power supply.

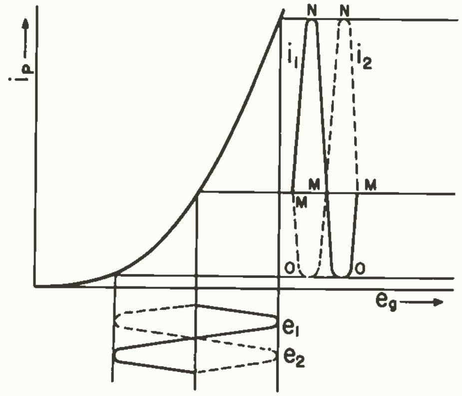


Fig. 15. When the operating point is on the lower bend of the tube characteristic curve, the plate current pulses i_1 and i_2 will be distorted.

Distortion Cancellation.

We mentioned previously that second-harmonic distortion generated within a push-pull amplifier is cancelled in the stage. Let's see how this distortion is produced and how it is cancelled.

First, we'll consider the beam power circuit shown in Fig. 13. Remember that the characteristic curve of a tube is not a straight line -- it is curved something like the line shown in Fig. 15. If the tube is operated on the bent portion of the curve shown in Fig. 15, distortion will be introduced.

Referring back to Fig. 13, let's consider the input signal across the primary of T_1 as "e" and in Fig. 15 consider the signals across the two halves of the secondary as e_1 and e_2. The plate current flowing through V_1 will be referred to as i_1 and the plate current flowing through V_2 as i_2. Let's look first at the signal e_1

applied to the tube V_1 -- this is represented in Fig. 15 as the solid line. When e_1 swings in a positive direction, i_1 increases as shown; when e_1 swings in a negative direction, then i_1 decreases as shown.

Notice, however, that the waveshape for the plate current is not equal in amplitude on the two sides of the zero axis -- that is, the alternation M-N-M is greater than the alternation M-O-M. This indicates that the stage has added second harmonics as well as other even harmonics to the original fundamental sine wave. In other words, it has added distortion (signals that were not present in the original signal).

At the same time, the other tube is getting the grid signal e_2, represented by the dotted line in Fig. 15. Notice also that the plate current i_2 is distorted on the M-O-M alternation. However, the plate current alternation of i_1 that is distorted occurs at the same moment as the portion of i_2 that is not distorted, and vice versa. This is due to the fact that e_1 and e_2 are 180° out-of-phase (when one is going positive, the other is going negative).

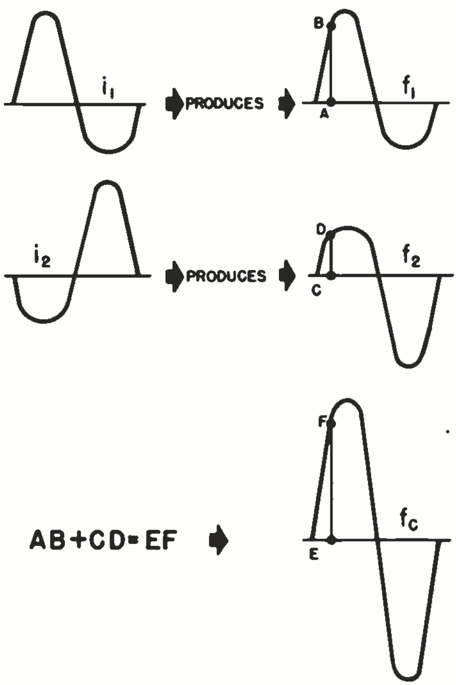


Fig. 16. The even-harmonic distortion cancellation occurs because of the manner in which the fluxes add in the output transformer.

In a transistor circuit such as is shown in Fig. 14, essentially the same thing happens. When the collector current of Q_1 is increasing, the collector current in Q_2 is decreasing. Similarly, when the collector current in Q_1 is decreasing, the collector current in Q_2 is increasing. Thus, if we have distortion in one-half of the output waveform produced by either transistor, we have exactly the same thing we have in the vacuum-tube stage, represented graphically in Fig. 15.

Now let's examine the action occurring in T_2 when the plate current of the two tubes (shown in Fig. 13) or the collector current of the two transistors (shown in Fig. 14) flows through the primary winding. This is shown in Fig. 16. Although the currents i_1 and i_2 are 180° out-of-phase, they now flow in opposite directions through the two halves of the primary T_2.

If i_1 produces flux f_1 and i_2 produces flux f_2 (as shown in Fig. 16), the fact that i_2 is flowing through the transformer in a direction opposite from that of i_1 means that flux f_2 will add to flux f_1, as if the two were being produced by a single current flowing through the entire primary winding. When f_1 and f_2 (in Fig. 16) are added, the resultant flux (f_c) is not distorted.

Because the resultant flux f_c is

not distorted even though the flux produced by the individual tubes or transistors is distorted, a push-pull amplifier will cancel out all even harmonics produced within the stage. In other words, if some second and fourth harmonic distortion is produced within the stage, it will be cancelled out in the stage because of the way in which the signals are recombined in the output transformer. This applies only to even-order harmonics produced within the stage because of a non-linear tube or transistor characteristic. A push-pull amplifier will not cancel even-order harmonics fed to it in the input, nor will it cancel out odd-order harmonics such as the third and fifth harmonics generated within the stage.

Push-pull amplifiers are not entirely distortion-free because a certain amount of third, fifth and higher odd harmonics will be produced within the stage. These harmonics are not cancelled out by another stage. Therefore, the amount of third harmonic distortion in the stage is usually what limits the amount of power we can get out of the stage without excessive distortion. The more power we try to get out of the stage, chances are the more third harmonic distortion there will be produced.

CLASS B AMPLIFIERS

A Class B push-pull amplifier using beam power tubes is shown in Fig. 17. Notice that this circuit is almost identical to the circuit shown in Fig. 13. The only difference is that the bias for the tubes is obtained from a separate source rather than from a cathode bias resistor. In Class B operation the tubes are biased approximately at cut-off. You cannot have current flow in the cathode circuit to develop this bias if the tubes are at cut-off and therefore grid bias for the tubes must be obtained from another source. It is usually obtained by placing a resistor in the negative side of the power supply.

In the Class B amplifier with the tubes biased to cut-off, when the in-

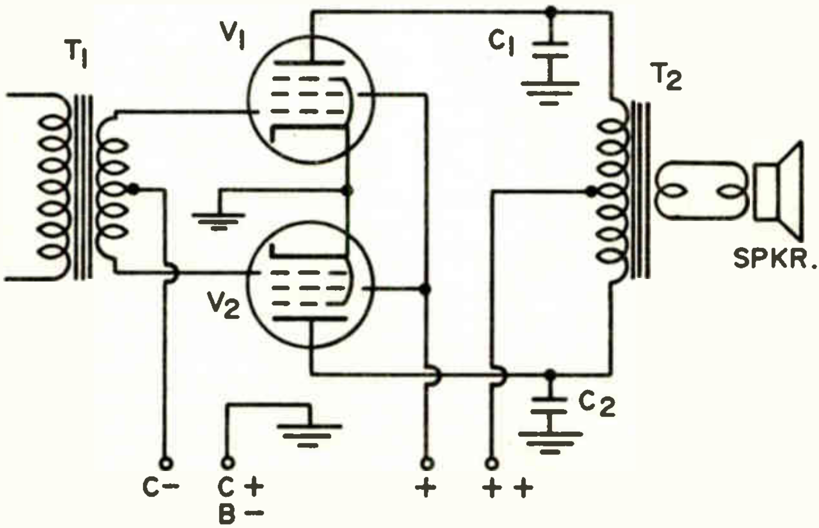


Fig. 17. Class B push-pull beam power tubes.

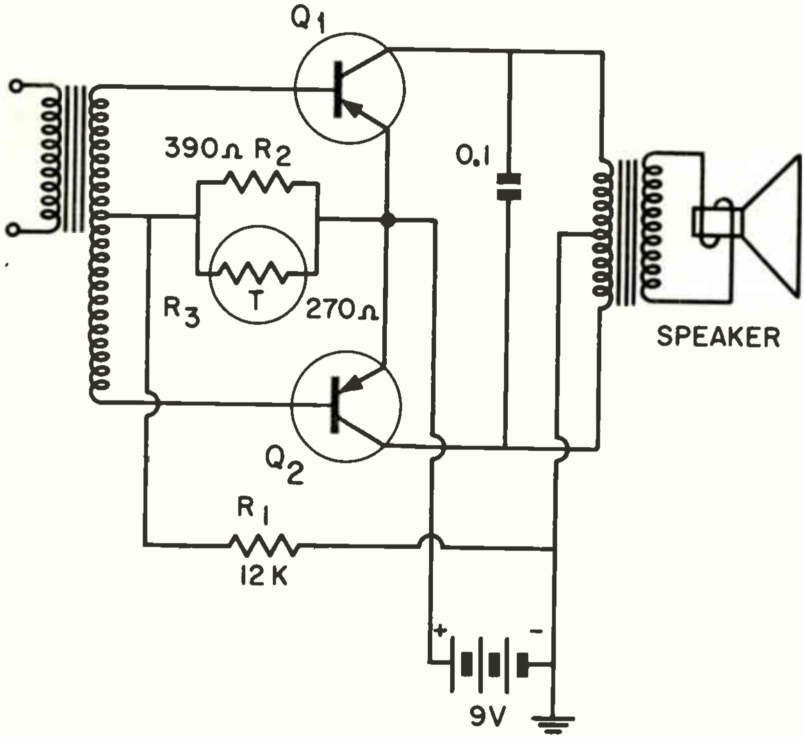


Fig. 18. Class B push-pull PNP transistors.

put signal swings the V_1 grid in a positive direction the tube will conduct. At the same time, the grid of V_2 will be driven in a negative direction so that no current will flow through V_2. During the next half-cycle, the opposite happens; the grid of V_2 is driven in a positive direction so that this tube conducts and no current flows through V_1.

Notice that we have indicated the plate and screen going to different voltage sources. As in the circuit shown in Fig. 13, the screen is returned to the output filter capacitor and the plate to the input filter capacitor in order to get the higher voltage on the plate.

Class B push-pull amplifiers using PNP transistors are shown in Fig. 18. Notice that the circuit is not too different from the tube circuit. The transistors are used in a common-emitter circuit. Bias is obtained by means of a voltage divider consisting of R_1 in series with the parallel combination of R_2 and R_3. The divider network is connected across the battery. When the input signal drives the base of Q_1 in a negative direction, Q_1 will conduct quite heavily. Holes will cross the emitter-base junction and flow over to the collector. At the same time as the base of Q_1 is driven in a negative direction, the base of Q_2 will be driven in a positive direction so that the base will be positive with respect to the emitter. Thus there will be reverse bias across the emitter-base junction of this transistor, and no current will flow through it. During the next half-cycle when the signal is reversed, the base of Q_1 will be driven in a positive direction so that a reverse bias will be placed across the emitter-base junction of this transistor and hence no current will flow through it. The base of Q_2 will be driven in a negative direction, increasing the forward bias so that the current flow through the transistor is increased. The number of holes crossing the emitter-base junction increases, and therefore the number of holes reaching the collector will increase.

The resistor R_3 in the emitter circuit is for bias stabilization purposes. This resistor is a thermistor. If the temperature is high, the resistance of R_3 will go down. This will reduce the forward bias across the emitter-base junction of the transistors and prevent the current from becoming excessive.

CLASS AB AMPLIFIERS

Quite frequently, vacuum tubes are not operated as Class B power amplifiers, but rather as Class AB power amplifiers. This is particularly true of amplifiers using beam power tubes. You can obtain almost as much power out of a Class AB vacuum tube as you can from the Class B circuit and the distortion is usually somewhat less with the Class AB operation. You will remember that for Class AB operation, bias on the tubes is about halfway between the bias used for Class A operation and the bias used for Class B operation. If the grids are driven positive at any time during the input cycle, we refer to the operation as Class AB_2 operation. If the grids are not driven positive then it is referred to as Class AB_1 operation.

When the power amplifier is operated as a Class AB_2 amplifier or a Class B amplifier where the grids of the tube are actually driven positive, the tubes consume power in the grid circuit. In an application of this type, the input transformer must be a step-down transformer in order to provide the power needed in the grid circuit. In addition, instead of using a voltage amplifier in the preceding stage, a power amplifier stage is required. In other words, in a circuit such as is shown in Fig. 17, the stage driving the push-pull Class B power amplifiers would probably be a single-ended Class A power amplifier. The Class A power amplifier produces the power required in the grid circuit of the two tubes. In Class A or Class AB_1 operation where the grids are not driven positive, the tubes do not require grid power and can be driven by a voltage amplifier stage.

SELF-TEST QUESTIONS

(u) In Fig. 13, why is the cathode resistor R_1 not bypassed?

(v) In Fig. 13, what purpose do C_1 and C_2 serve?

(w) In Fig. 14, is T_1 a step-up or a step-down transformer?

(x) What is one of the primary advantages of a push-pull type of power amplifier circuit?

(y) What is the difference between the Class A power amplifier in Fig. 13 and the Class B power amplifiers in Fig. 17?

(z) In Fig. 18, how are the transistors biased?

Reducing Distortion

Distortion is one of the most serious problems in low-frequency amplifiers. There are three kinds of distortion with which we must contend: amplitude distortion, frequency distortion, and inter-modulation distortion.

Amplitude Distortion.

Amplitude distortion results from the creation of irregularities in amplifying the signal. For example, one-half of a sine wave does not receive the same amplification as the other half, and harmonic distortion is produced. This means that signals that are multiples of the original signal frequency are produced. If the signal has twice the frequency of the original signal, it is called second harmonic distortion; if the signal is three times the frequency of the original signal, it is called a third harmonic distortion, and so on.

A small amount of distortion of this type is hardly noticeable, but the quality of the amplified signal usually suffers if you get a certain amount of second harmonic distortion plus some third (and higher) harmonic distortion.

Sometimes, in amplifying a sine wave, a small pip or irregularity may appear on the sine wave that was not present in the original signal. This type of distortion is also called amplitude distortion.

Frequency Distortion.

You already know that the gain of an amplifier falls off at low frequencies due to the reactance of the coupling capacitor used between the stages of an amplifier. Thus, low-frequency signals in some amplifiers do not receive the same amount of amplification as the middle-frequency signals receive. Similarly, at some high frequency, the various capacities in the circuit begin to reduce the gain of the amplifier at high frequencies so that high-frequency signals do not receive the same amount of amplification as middle-frequency signals receive. This failure to amplify signals at all frequencies equally is known as frequency distortion.

Inter-Modulation Distortion.

If an amplifier is not operated on the linear portion of its characteristic curve, mixing of two or more signals in the amplifier sometimes occurs. For example, suppose that two signals are fed to an amplifier at the same time. This might happen when an amplifier is amplifying musical notes of two different frequencies. If the two signal frequencies beat together to produce new signals equal to the sum and difference of the original signal frequencies, we have what is called inter-modulation distortion. In this case, the two signals have mixed together.

Inter-modulation distortion can be kept to a minimum by operating the amplifier on the linear portion of its characteristic curve and by avoiding operation of the amplifier at or near its maximum amplification capabilities.

There are several methods used to keep distortion to a minimum. One of these, of course, is careful design; another is making sure that the tube or transistor used in the amplifier is operated as it should be. This, however, is not enough to get what we would call good high-fidelity reproduction, although it is good enough for most table-model radio receivers (where we do not expect extremely high-quality output) and for television receivers (where we are primarily interested in sound to accompany the picture). In quality sound-reproducing equipment (high-fidelity or stereo, for example), on the other hand, steps must be taken to keep distortion as low as possible. One of the simplest steps in reducing distortion is the use of inverse feedback.

INVERSE FEEDBACK

Fig. 19 illustrates two simple examples of inverse feedback. The circuit at A shows a vacuum tube; the circuit at B, a transistor.

Looking first at the vacuum-tube circuit, we see that it is exactly the same as the circuit you studied previously, but in this case the bias resistor (R_2) in the cathode of the circuit of the tube is not bypassed.

The purpose of the cathode bypass capacitor is to bypass signals around the cathode resistor, R_2. We pointed out previously that all capacitors have a certain reactance, and that at low frequencies this capacitor may not be a good bypass. As a result, there will be some attenuation of low-frequency signals. By completely eliminating the capacitor, we can eliminate the problem. Now all signals will receive some reduction in amplification because there is a

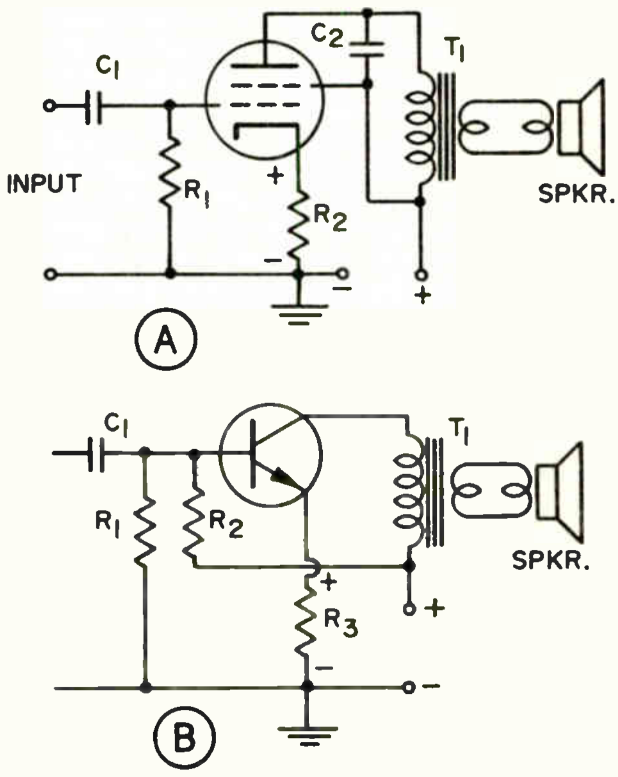


Fig. 19. Simple inverse feedback circuits.

certain amount of feedback in the circuit.

Let's consider what happens when a signal drives the grid of a tube in a positive direction. This causes the current flow through the tube to increase; hence the voltage drop across R_2 (which has the polarity shown on the diagram) increases. Thus the cathode swings in a positive direction. The signal voltage developed across R_2 is subtracted from the grid voltage. We then have a certain amount of feedback; in other words, the signal from the output circuit is fed back into the input circuit. We call this feedback "inverse" because the signal subtracts from the input signal.

In the transistor circuit shown in Fig. 19B, the resistor R_3 is put in the emitter circuit for temperature stabilization. It is usually bypassed, but the bypass capacitor acts in exactly the same way as the cathode-bypass capacitor in the tube circuit. At low frequencies, the capacitor is not an effective bypass. By eliminating the capacitor a certain amount of inverse feedback is introduced into the circuit. When the input signal drives the base in a positive direction, the forward bias across the emitter-base junction increases and the current flow through the transistor increases. This causes the voltage across R_3 to increase and the increase in voltage subtracts from the positive base voltage, thus reducing the net base-emitter signal voltage.

The feedback signal will be developed at all signal frequencies and will tend to make the gain of the amplifier more constant over a wider frequency range. You can see why this is so if you consider that at some frequency the signal reaching the amplifier has a higher amplitude than at other frequencies. This signal in the case of the circuit at A will cause a higher cathode current through the tube. Hence the voltage across R_2 will be higher than at other frequencies and will tend to reduce the input signal more than at other frequencies. Similarly, if the signal applied to the base is greater at some particular frequency, it will tend to develop a higher feedback signal across R_3. This, in turn, will tend to reduce it more, keeping it closer to the amplitude of the other signal frequencies.

We sometimes call this type of feedback "degeneration". Degeneration and inverse feedback are essentially the same thing. When the signal fed from the output circuit back to the input circuit reduces the input signal, we say this is degenerative feedback or inverse feedback. When the signal fed from the output circuit back to the input circuit adds to or aids the input signal, we call it regenerative feedback. Regenerative feedback is never used to improve the response of an amplifier.

The example of inverse feedback shown in Fig. 19 is in both cases contained within a single stage. Inverse feedback can be used over more than one stage as shown in the example in Fig. 20. Here a signal is fed from the plate of the output tube back to the cathode of the voltage amplifier tube.

In the circuit shown in Fig. 20, when the input signal drives the grid of V_1 in a positive direction, a negative-going signal will be developed in the plate circuit. This signal is fed to the grid of V_2 through the capacitor C_3 and this in turn causes a positive-going signal in the output circuit of V_2. The positive-going

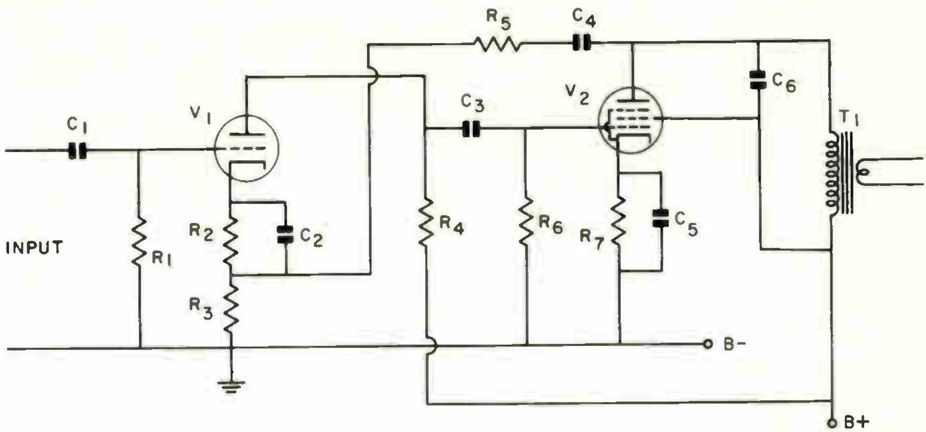


Fig. 20. Inverse feedback from the plate of the power output tube to the cathode of the voltage amplifier tube.

signal is fed through C_4 and R_5 back into the cathode circuit of V_1 where it subtracts from the input signal fed to V_1. The advantage of two-stage feedback of this type is that it tends to equalize the gain of the amplifier over both stages rather than in just a single stage as in the circuits shown in Fig. 19.

PHASE INVERTERS

One of the causes of distortion in push-pull power amplifiers (such as in Fig. 13) is the input transformer. Transformers that respond equally to a wide range of frequencies are difficult and expensive to manufacture. The transformer in this circuit can be eliminated by means of a phase inverter stage as shown in Fig. 21.

In this circuit the input signal is applied between the grid and cathode of V_{1A}. It is amplified by this tube and the signal phase is inverted. The signal is fed through C_3 to the grid

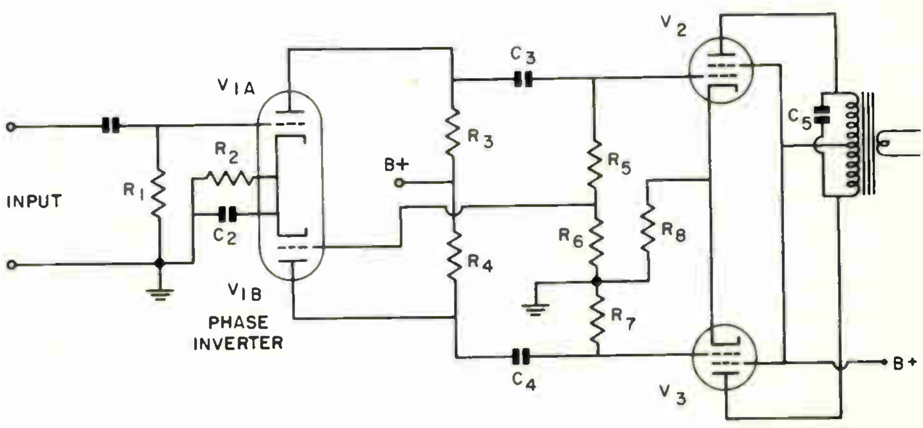


Fig. 21. A typical circuit using a phase inverter to drive one stage of a push-pull amplifier.

of V_2. Meanwhile, the amplified signal is divided by the resistors R_5 and R_6. Usually, R_5 is considerably larger than R_6 so that only a small part of the amplified signal is taken and fed back to the grid of V_{1B}, the phase inverter stage. Remember that this signal will be in phase with the signal fed to the grid of V_2. The signal is amplified by V_{1B}; its phase is inverted and then fed through C_4 to the grid of V_3. Thus V_2 and V_3 are driven by signals 180° out-of-phase.

If the ratio of R5 and R6 are selected correctly, the amplified signal fed to the grid of V_3 will be equal to the amplified signal fed to the grid of V_2. This type of circuit provides the two signals equal and 180° out-of-phase to drive the push-pull power amplifier. At the same time, better frequency response can be obtained with this type of circuit than with the circuit shown in Fig. 13 using an input transformer.

A circuit similar to the one shown in Fig. 21 could be used with transistor push-pull stages, but here the input transformer is primarily an impedance-matching device and usually there is more to be gained by using the input transformer than by attempting to use the phase-inverter type of stage found in push-pull vacuum-tube amplifiers. There are other things that can be done with transistor circuits to get improved results that cannot be done with vacuum-tube circuits. We will look at these circuits now.

TRANSISTOR CIRCUITS

A two-transistor amplifier which uses a PNP transistor and an NPN transistor is shown in Fig. 22. This amplifier makes use of the characteristics of the two different transistors to eliminate the coupling capacitor usually found between amplifier stages, and in so doing also eliminates distortion due to the capacitor.

In the circuit shown, the input impedance is controlled at high frequencies by R_5. C_1 and C_2 have a low reactance at middle and high frequencies, so the total impedance across the input at high frequencies

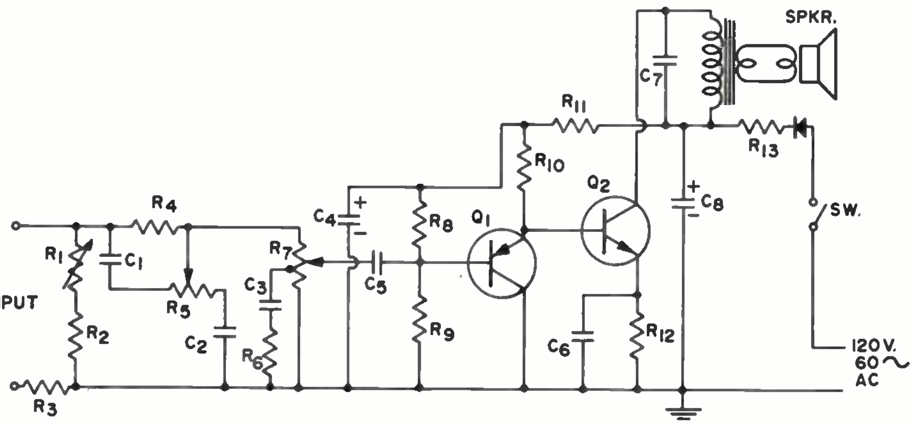


Fig. 22. Two-stage transistor amplifier.

will be equal to the resistance of R_5. At low frequencies, R_1 varies the input impedance. By varying the setting of R_1, the input impedance (and hence the amplitude of any low-frequency input signal) can be varied. Thus R_1 serves as a bass or low-frequency tone control. R_4 is in the input circuit between the input and the volume control. It can be bypassed at high frequencies by the setting of R_5. When the sliding contact is up towards C_1, C_1 provides an effective bypass around R_4 at high frequencies. When the sliding contact is down at the other end, R_4 and R_5 are essentially in parallel at high frequencies, and this tends to reduce the amplitude of the high-frequency signals. Thus, R_5 serves as a high-frequency tone control, which is usually called a treble control.

R_7 is the volume control. Notice that this control has a tap on it, and that C_3 and R_6 are connected between the tap and ground. This type of circuit is referred to as automatic bass compensation. When you turn the volume control to the low-volume position, the low-frequency sounds appear weaker than the high-frequency sounds. The higher-frequency signals are attenuated by means of C_3 and R_6, so that there is a tendency to equalize the loudness of the high-frequency and low-frequency signals. Actually, the low-frequency signals are given greater amplification to compensate for the fact that they are less noticeable at low volume levels. The transistor Q_1 is used in a common collector circuit. The input signal is fed to the base through C_5. The output is taken across the emitter resistor (R_{10}) and is fed directly to the base of Q_2. Q_2 is a power transistor which is used in a common-emitter circuit.

The two-stage amplifier is designed to operate directly from a 120-volt power line. Notice that a half-wave rectifier circuit using a silicon rectifier is shown. Resistor R_{13} is a series resistor to limit the charging current through the rectifier when the equipment is turned on. C_8 is the input filter capacitor; C_4, the output filter capacitor. R_{11} is the filter resistor. The supply voltage applied to the collector of Q_2 is not filtered as well as the voltage fed to Q_1. The voltage fed to Q_2 does not require the filtering because if there is hum voltage present with the dc, it does not receive any amplification. However, any hum voltage on the dc applied to Q_1 will be amplified and will result in an objectionable hum in the output.

Another transistor circuit of interest is the one shown in Fig. 23. This circuit is referred to as a complementary-symmetry push-pull amplifier. Here an NPN transistor and a PNP transistor are used. When the input signal drives the base of Q_1 in a positive direction, the current through Q_1 will decrease because this transistor is a PNP transistor and the positive voltage applied to the base of the transistor tends to decrease the forward bias applied to it. At the same time, the positive input signal is fed to the base of Q_2 and this increases the emitter-base forward bias and causes the current through this transistor to increase. This current is represented by i_2 and flows through the output transformer primary in the direction shown. In the next half-cycle, when the signal swings negative, this will subtract from the forward bias across Q_2, reducing the

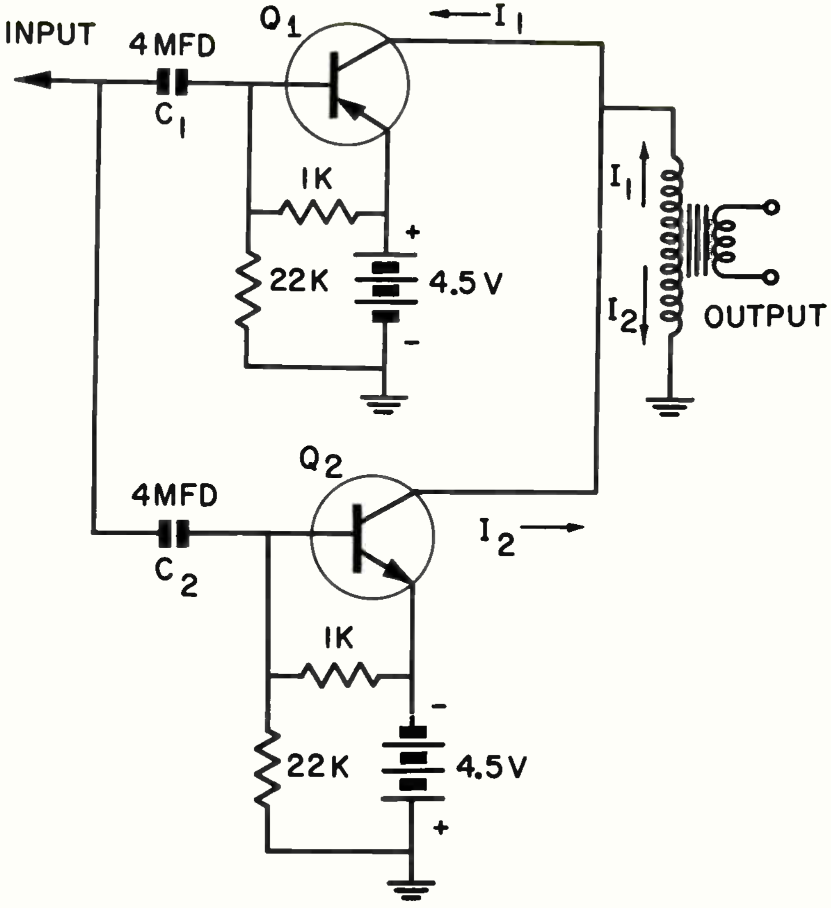


Fig. 23. Complementary symmetry push-pull amplifier.

current flow through it, and increase the forward bias across the emitter-base of Q_1, causing the hole flow through it to increase. The increase in hole flow through the transistor results in an increase in current flow i_1 through the primary of the output transformer in the direction shown.

If the transistors shown in Fig. 23 are balanced and biased essentially so that they are cut off without any signal flow, there will be no current flow through the primary of the transformer until a signal is applied. Then the current will flow through the transformer; the direction of flow will depend upon whether the signal is positive-going or negative-going. During one half-cycle the current will flow through the primary of the output transformer in one direction and during the other half-cycle it will flow through the primary of the transformer in the opposite direction.

Notice that this type of circuit eliminates the input transformer,

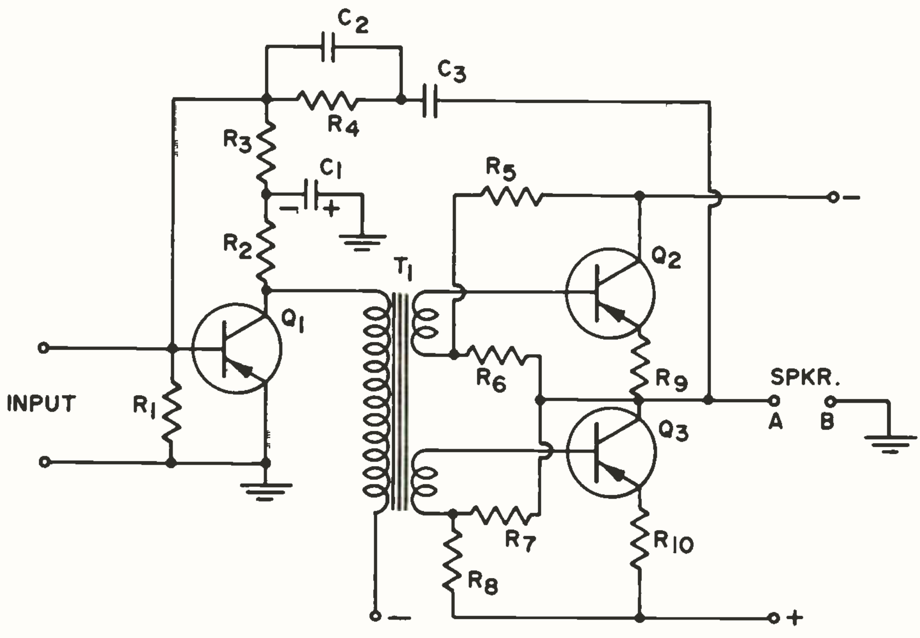


Fig. 24. Three transistor amplifiers where there is no output transformer.

yet we still have a push-pull power amplifier. In a circuit designed for power line operation, batteries would not be used and a single power supply capable of supplying the required positive and negative voltages would be used.

Another transistor circuit of interest is shown in Fig. 24. In this circuit the need for an output transformer has been eliminated. Notice that a negative voltage has been supplied to the collector of Q_2 and a positive voltage to the emitter of Q_3 through R_{10}. The negative voltage has the same value as the positive voltage with respect to ground. The resistors R_5 plus R_6 are equal in value to resistors R_7 and R_8 so that the junction of R_6 and R_7 will be at ground potential. Similarly, R_{10} is equal to R_9, and Q_2 and Q_3 are similar transistors so that there is an equal voltage drop across the emitter resistor and transistor in each case. The junction of the collector of Q_3 and resistor R_9 are also at ground potential. The junction of R_6 and R_7 and the collector of Q_3 and R_9 are connected together at terminal A. The speaker is connected between terminals A and B.

With no signal the transistors are biased essentially at cut-off so that there is little or no current flow through the transistors or through the speaker. When a signal is applied to the input circuit, it is amplified by Q_1 and fed to the transformer T_1. For example, suppose that the input signal drives the base of Q_1 in a positive direction. Since the transistor is a PNP transistor, this will reduce the forward bias across the emitter-base junction. This means that the number of holes reaching the collector will decrease and hence the current flowing through the primary of T_1 will decrease. The two secondary windings on T_1 are phased so that when this happens the voltage applied to the base of Q_3 will swing in a positive direction and the voltage applied to the base of Q_2 will swing in a negative direction. If Q_3 is already biased close to cutoff the positive voltage on the base will reduce the emitter-base junction forward bias still further so that if there is any current flowing through this transistor, it will drop even lower. On the other hand, the negative voltage applied to the base of Q_2 will increase the forward bias across the emitter-base junction. This will cause the number of holes crossing the junction to increase; the number of electrons flowing from the emitter of Q_2 through R_9 to terminal A and through the speaker to ground will also increase.

The current flowing through the speaker will develop a voltage drop across it so that the end connected to terminal A will be negative. This negative signal voltage will be fed through C_3 and the parallel combination of C_2 and R_4 back to the base of Q_1, where it will subtract from the input signal. Thus, this is an inverse feedback path.

When the signal swings in the opposite direction and drives the base of Q_1 in a negative direction the number of holes crossing the emitter-base junction will increase and hence the electron current through the primary of T_1 will increase. This will cause the base of Q_3 to swing in a negative direction. When this happens the number of holes crossing the emitter-base junction of Q_3 must increase. Therefore the number of electrons flowing from terminal B through the speaker to terminal A to the collector of Q_3 must increase. In flowing through the speaker in this direction, the electrons will develop a voltage so that terminal A is positive. This voltage is fed back to the input of Q_1 and will once again subtract from the input voltage.

The circuit is interesting in that the output transformer has been eliminated. The output transformer is one of the primary causes of distortion in low-frequency power amplifiers. Usually, there is more distortion developed in the output transformer than in the input transformer, because dc current flowing through the output transformer will be much higher than the dc current flowing through the input transformer. Also, there is more of a tendency to produce core saturation, which in turn produces more nonlinearity in the output transformer than in an input transformer.

A circuit of the type shown in Fig. 24 is not possible with vacuum tubes because of the high output impedance encountered in tubes. However, transistors are comparatively low-impedance devices, which makes this type of circuit practical.

SUMMARY

The circuits shown in this section of the lesson are all practical circuits. They are the type of voltage and power amplifiers you are likely to encounter in radio and television receivers and in high-fidelity and stereo equipment which you will be called upon to service. You should be familiar with the operation of these circuits so that when you come across them you will know how they work and be able to proceed to the cause of the trouble in a logical manner.

SELF-TEST QUESTIONS

(aa) What is distortion?
(ab) Name the three types of distortion.
(ac) What is inverse feedback?
(ad) In the circuit shown in Fig. 19, how does omitting the cathode bypass capacitor introduce inverse feedback?
(ae) In the circuit shown in Fig. 21, what is the purpose of the phase inverter stage?
(af) In Fig. 22, what purpose does R_1 serve?
(ag) In the circuit shown in Fig. 22, what purpose does R_5 serve?
(ah) In the circuit shown in Fig. 22, why is there a tap on the volume control on R_7? What purpose does this tap and the associated components serve?
(ai) In the circuit shown in Fig. 22, in what type of circuit is Q_1 used?
(aj) What is the primary advantage of the circuit shown in Fig. 24?
(ak) In Fig. 24, what purpose does the circuit serve which connects terminal A through C_3 and the parallel combination of C_2 and R_4 back to the base of Q_1?

ANSWERS TO SELF-TEST QUESTIONS

(a) R_1 is placed in the cathode circuit of V_1 to develop an operating bias for the tube. Current flowing from B minus through R_1 will develop a voltage having the polarity shown on the diagram. This makes the cathode positive with respect to ground. The grid of V_1 connects to ground through the generator and, since there will be no dc current flow through the grid circuit, there will be no voltage drop; therefore, the grid will be at dc ground potential. This will make the cathode positive with respect to the grid. In other words, the grid is negative with respect to the cathode.

(b) C_1 is a cathode bypass capacitor. Its purpose is to provide a low impedance path for the signal around the cathode bias resistor R_1. The value of C_1 is selected so that its reactance is much lower than the resistance of R_1 at the frequency of the signals to be amplified. Often, C_1 is an electrolytic capacitor; a large capacitor is required in this circuit because the resistance of R_1 is usually small and a large capacitor with a low reactance is required to bypass it effectively.

(c) In Fig. 1, the amplified signal voltage developed across R_2 will be 180° out-of-phase with the input signal voltage.

(d) No. The voltage drop across R_1 has the wrong polarity to provide a forward bias across the emitter-base junction of Q_1. The forward bias across the emitter-base junction is provided by the battery. The small current flowing through R_1 develops a voltage across this resistor having the polarity shown. This voltage subtracts from the battery voltage so that the net forward bias across the emitter-base junction is equal to the battery voltage less the voltage drop across R_1. Usually, the forward bias of only a few tenths of a volt is required across the emitter-base junction and R_1 is selected so that the base current flowing through it will develop a voltage drop, and when it is subtracted from the battery voltage the forward bias will be only a few tenths of a volt.

(e) No. The holes flow to the collector, where they are filled by electrons. These electrons come from the negative terminal of the battery and flow through R_2 to the collector. The movement of the holes is contained entirely within the transistor. In the external circuits, all current flow is electron flow.

(f) The forward bias placed across Q_1 makes the base negative and the emitter positive. When the input signal tends to swing the base in a positive direction it reduces the forward bias across the emitter-base junction. When this happens the number of holes crossing the junction decreases. When the number of holes crossing the emitter-base junction goes down, the number of holes reaching the collector will go down. This means that fewer electrons will flow through R_2 to fill the holes reaching the collector. When the number of electrons flowing through R_2 decreases, the voltage drop across R_2 will go down, so that the polarity of the collector end of R_2 will swing in a negative direction. In other words, the negative voltage on the collector of Q_1 will increase.

(g) In the circuit shown in Fig. 2, the value of C_3 is selected so that it will have a very low reactance over the frequency range of the signals to be amplified. Therefore, as far as the signal is concerned C_3 acts as a short circuit. In other words, the circuit functions as though there is a direct connection from both ends of R_2 to the end of R_3. The input resistance of Q_2 is equal to R_3 in parallel with the resistance across the emitter-base junction. The resistance across the emitter-base junction is comparatively low; therefore, the input resistance of Q_2 is low. The signal voltage developed in the collector circuit of Q_1 will depend upon the size of the load resistor. Thus, even though R_2 may be made large in order to develop a fairly high signal voltage, since the input resistance of the second stage is directly in parallel with this resistor it will pull the effective value of the collector load resistance down. As the net result, the gain of Q_1 is affected appreciably by the input resistance of the second stage, Q_2.

(h) The coupling capacitor C_3.

(i) The high frequency limit of the amplifiers is controlled primarily by the capacity in the output of the first stage and the input of the second stage. In Fig. 1 the output capacity is made up of the plate-to-ground capacity of V_1 plus wiring capacity and the grid-to-cathode capacity of V_2 plus wiring capacity. In Q_1 the output capacity of the transistor plus wiring capacity limits the frequency response along with the input capacity of Q_2 plus wiring capacity in this circuit.

(j) The early vacuum tubes were not capable of giving a very high gain. By means of a step-up transformer between stages the gain obtainable from the stage could be increased by the turns ratio of the transformer.

(k) The high gain available with modern vacuum tubes makes transformer coupling no longer necessary. In addition, R-C coupling is more economical and yields better frequency response.

(l) The transformer serves as an impedance-matching device. It matches the high output impedance of the first stage to the low input impedance of the second stage. Thus, maximum power transfer is possible and a higher voltage gain can be obtained.

(m) Modern transistors are relatively inexpensive. By using a three-stage R-C coupled voltage amplifier you can obtain as much gain as with a two-stage transformer coupled amplifier. The R-C coupled amplifier will be less expensive and also capable of better frequency response.

(n) You must obtain a replacement transformer having the same turns-ratio as the turns-ratio of the original transformer.

(o) The power amplifier in a radio receiver is designed to supply the power required to drive or operate the loudspeaker from the signal voltage applied to it.

(p) A beam power tube or a pentode tube. Triode tubes have been used as power amplifiers but they require considerable driving power. Therefore, it is not likely you will find a triode tube in modern equipment.

(q) The bypass capacitor in the plate circuit of a power amplifier is used to bypass the high-frequency audio signals and prevent high-frequency oscillation.

(r) Low-frequency distortion can be kept to a minimum by using a large value capacitor for C_2, the cathode bypass capacitor. This will reduce the frequency at which the reactance of the capacitor becomes large enough to stop acting as an effective bypass.

(s) The collector which is connected to the case of the transistor must dissipate a fair amount of heat. Since the collector is operated at almost ground potential, very little insulation is required between the collector and ground. Thus a good heat contact can be made between the collector and receiver chassis so that the chassis can aid in dissipating the heat from the transistor.

(t) You are more likely to find a push-pull output stage where the transistors are operated in a Class B circuit.

(u) The cathode bias resistor is not bypassed because the current through it remains essentially constant. Any increase in current through V_1 is compensated for by an equal decrease in current through V_2 and vice versa.

(v) C_1 and C_2 are plate bypass capacitors. They prevent high-frequency oscillation.

(w) T_1 is a step-down transformer. A step-down transformer is required to match the low input impedance of Q_1 and Q_2 to the high output impedance of the driver stage.

(x) In a push-pull power amplifier, even-order harmonics generated within the stage are cancelled within the stage so that the distortion is kept quite low.

(y) The circuits are practically identical except that a higher bias is used in Fig. 17 in order to bias the tube essentially at plate current cut-off. In addition, T_1 will be a step-down transformer in the circuit shown in Fig. 17, although it can be a step-up transformer in the circuit shown in Fig. 13.

(z) Forward-bias across the emitter-base junctions in the circuit shown in Fig. 18 is provided by the voltage divider consisting of R_1 in series with the parallel combination of R_2 and R_3. This will bias the base slightly negative with respect to the emitter.

(aa) Distortion is the introduction of a signal in the output that is not present in the input or the loss of a signal in the output that is present in the input.

(ab) Amplitude distortion, frequency distortion and intermodulation distortion.

(ac) Inverse feedback is a signal fed from the output of a circuit back to the input, with a polarity such that it subtracts from the input signal. With inverse feedback, the feedback signal reduces the amplitude of the input signal.

(ad) With the cathode bypass capacitor omitted, the signal current flowing through the tube will develop a voltage across R_2. Since this voltage has the same polarity as the grid voltage producing it, it reduces the net grid-to-cathode voltage. The voltage thus subtracts from the input voltage.

(ae) The phase inverter eliminates the need for an input transformer to the push-pull output stage. It takes part of the signal fed to V_2, one of the output stages, and inverts it so that it can be used to drive V_3.

(af) R_1 serves as a low-frequency tone control. This type of control is called a base control.

(ag) R_5 serves as a high-frequency tone control. It is usually called a treble control.

(ah) The tap on the volume control along with capacitor C_3 and R_6 form an automatic bass compensation circuit. At low volume levels, low-frequency sounds sound weaker than high-frequency sounds. To balance the relative loudness between the two, the high-frequency signals are bypassed through C_3 and R_6 so that the low-frequency signals receive greater amplification.

(ai) Q_1 is used in a common-collector circuit.

(aj) The elimination of the output transformer.

(ak) This circuit provides inverse feedback to improve the frequency response of the amplifier.

Lesson Questions

Be sure to number your Answer Sheet B202.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. When considering the action of the coupling network at middle frequencies, why can we ignore the coupling capacitor C_3 in Figs. 1 and 2?

2. In Fig. 2, what percentage of the signal voltage appearing across R_2 will appear across R_3 at the low-frequency half-power point?

3. Why is more gain obtained from the two-stage transformer-coupled voltage amplifier shown in Fig. 9 than from the two-stage voltage amplifier circuit shown in Fig. 2?

4. If transformer coupling between transistor voltage amplifiers provides maximum gain, why is R-C coupling more frequently used?

5. What is the difference between a voltage amplifier and a power amplifier?

6. In Fig. 14, will the voltage drop across R_2 add to or subtract from the forward bias across the emitter-base junction of Q_1 and Q_2?

7. Why are push-pull Class B power amplifiers preferred over push-pull Class A power amplifiers or single-ended power amplifiers in transistor portable receivers?

8. What is the advantage of using inverse feedback in low-frequency voltage and power amplifier stages?

9. In the two-stage amplifier shown in Fig. 22, in what type of circuit are the transistors Q_1 and Q_2 used?

10. What will be the polarity of the voltage at terminal A, in the circuit in Fig. 24, when the input signal drives the base of Q_2 in a positive direction and the base of Q_3 in a negative direction?

CONVERSATIONALLY SPEAKING

Conversation is a give-and-take proposition, and listening is the "take" part. Talk only when you can say something of interest. Otherwise, remain silent. Let your silence be eloquent enough to show that you derive pleasure from listening--that you consider the words of your companion far more valuable than anything you could say. This kind of silence can make just as many friends as good conversation.

Talk about things which will interest and please your listeners. Their hobbies, their work, their children and their homes are all good opening topics for conversation. Don't talk about yourself, your troubles or your work, unless asked.

Avoid expressing definite opinions on controversial subjects, for they often lead to unpleasant arguments. Ridicule of another person is likewise taboo at all times. If you can't say pleasant things about others, keep quiet. Finally, reserve technical discussions for technically minded listeners.

nri

McGraw-Hill

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-215

ACHIEVEMENT THROUGH ELECTRONICS

nri

HOW RADIO-FREQUENCY AMPLIFIERS WORK

B203

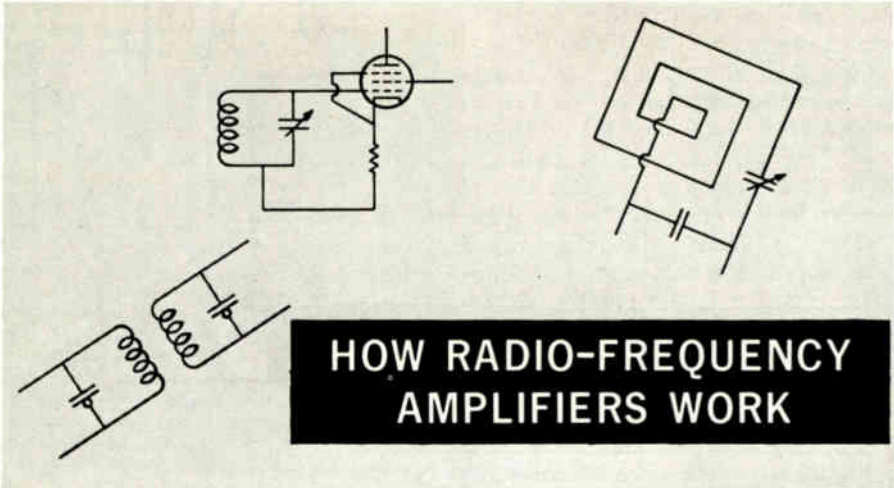
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

HOW RADIO-FREQUENCY AMPLIFIERS WORK

B203

STUDY SCHEDULE

JDL4M771 1975 EDITION Litho in U.S.A.



HOW RADIO-FREQUENCY AMPLIFIERS WORK

A radio-frequency amplifier is an amplifier designed to amplify signal frequencies above the sound or audio range. Some radio-frequency amplifiers are designed to operate at frequencies as low as 20 or 30 kilocycles; you will find amplifiers of this type in very low-frequency communications equipment. Other radio-frequency amplifiers are designed to operate at frequencies of several hundred megacycles. The radio-frequency amplifier found in the vhf tuner of a television receiver must be able to operate at frequencies over 200 megacycles in order to amplify signals from TV Channels 11, 12 and 13. As you will see in this lesson, the operation of low-frequency radio-frequency amplifiers is essentially the same as those designed to operate in the very high-frequency regions although some of the problems encountered in amplifiers operating at several hundred megacycles or more are not encountered in amplifiers operated at low frequencies.

Radio-frequency amplifiers are usually referred to as rf amplifiers. RF amplifiers are found in all types of electronic equipment. Modern radio and television receivers all contain rf amplifiers. RF amplifiers are used in all types of communications equipment designed to transmit or receive information by means of radio waves ranging from simple inexpensive portable receivers up to the complex radio systems designed to keep in touch with satellites and relay information about the earth or other planets from the satellite back to earth.

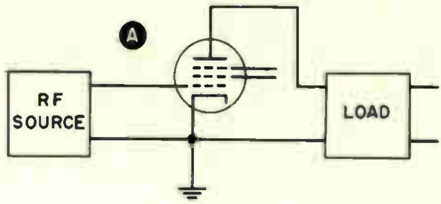
RF amplifiers perform two important functions. They amplify weak radio-frequency signals and they help select the desired signal while rejecting undesired signals.

The signal transmitted by a radio or television station is comparatively weak by the time it reaches the antenna of the radio or television receiver. Before the information contained in that signal can be extracted from it, the strength of the signal must be built up. Radio and television receivers have rf amplifiers in them to build up the weak signal picked up by the receiving antenna.

As you probably already know, in this country there are hundreds of radio stations sending out signals in the standard radio broadcast band; there are several hundred television stations sending out television signals; and in addition there are thousands of stations used for commercial communications. A radio or television receiver must be able to select the one signal you are interested in from all these and reject the others. The various radio broadcast stations in one locality all operate on different frequencies. Your receiver has rf stages in it that can be tuned to respond to one frequency while rejecting signals of other frequencies. Thus, in addition to amplifying the weak signals picked up by your receiving antenna the rf amplifier helps to select the one signal you want and reject the others.

In modern superheterodyne receivers you will remember that the incoming signal is mixed with the signal generated by a local oscillator (in the set) and a new signal frequency is produced. This frequency is called the intermediate (i-f) frequency. This i-f is actually an rf frequency and the amplifiers designed to amplify it are rf amplifiers. However, to distinguish between these amplifiers and the ones that amplify the incoming signal before its frequency is changed, they are usually called i-f amplifiers. Our discussion in this lesson will cover both the rf amplifiers that amplify the signal before the signal frequency is changed and the i-f amplifiers that amplify it after the signal frequency has been changed.

Because rf amplifiers are so widely used in both transmitting and receiving equipment, they are considered a basic circuit and all technicians should be familiar with and understand their operation. You will run into this type of amplifier regardless of what branch of the electronics field you work in. Before we go further into the subject let's see what a basic rf amplifier looks like.



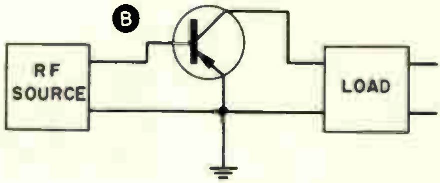


Fig. 1. Two basic rf amplifiers are shown above. Notice that they are similar to the low-frequency amplifiers you have already studied.

A BASIC RF AMPLIFIER

Two basic rf amplifiers are shown in Fig. 1. In Fig. 1A we have shown a circuit using a vacuum tube. We have not shown the screen-grid or suppressor-grid connections to the tube nor have we shown any bias on the tube, because they do not enter into our consideration of the basic rf stage. Notice that the rf source signal is applied between the grid and the cathode of the tube, and that there is a load in the plate circuit of the tube. The amplified rf signal voltage is developed across this load.

In the circuit shown in Fig. 1B we have shown a transistor using a common emitter circuit. Again, the rf signal source is applied between the base and the emitter, and the load is placed in the collector circuit of the transistor. The amplified signal voltage appears across the load.

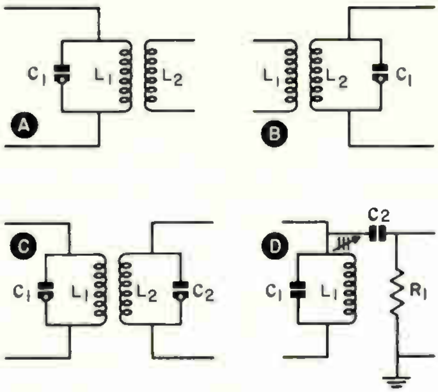


Fig. 2. Four basic loads which might be found in rf amplifiers.

At first glance it may appear that these circuits do not differ appreciably from the low-frequency amplifiers that you have already studied. However, there is a big difference, and that difference is in the load used in the output circuit of the rf amplifier. RF amplifiers almost always use some type of resonant circuit as the load.

Four basic loads which might be found in rf amplifiers are shown in Fig. 2. In the circuit shown in Fig. 2A we have an rf transformer. Coils L_1 and L_2 are inductively coupled together. The primary L_1 is tuned to resonance by the capacitor C_1. The parallel resonant circuit thus formed is connected into the plate circuit of the rf amplifier.

In Fig. 2B we have another type of rf transformer. Here L_1 and L_2 are again inductively coupled together. In this circuit, however, the secondary, L_2, is tuned to resonance by C_1. Thus, instead of having the resonant circuit in the output circuit of the rf amplifier we simply have L_1 in the output circuit, and the resonant circuit is in the input of the following rf stage.

In the circuit shown in Fig. 2C we have another type of rf transformer. Again, L_1 and L_2 are inductively coupled together, but in this circuit both L_1 and L_2 are tuned to resonance. L_1 is tuned to resonance by C_1, and L_2 is tuned to resonance by C_2. This type of rf transformer is found in the i-f amplifiers of most radio receivers and some TV receivers.

Another load that might be found, particularly in TV receivers, is shown in Fig. 2D. In this circuit C_1 is a fixed capacitor and L_1 a variable inductance. The circuit is tuned to resonance by adjusting a slug which moves in and out of L_1 to vary the inductance of the coil. This method of varying the inductance in the circuit rather than the capacity could also be used in circuits like those shown in Figs. 2A, 2B, and 2C. L_1 and C_1 form a parallel resonant circuit. The signal appearing across this parallel resonant circuit is coupled to the following stage through capacitor C_2. In some high-frequency circuits, C_1 in Fig. 2D is omitted. The output capacity of the tube or transistor and the wiring capacity in the circuit take the place of C_1.

This brief look at the basic rf stage and the loads that you are likely to find in the plate circuit of this type of stage should immediately point out to you the importance of resonant circuits in rf amplifiers. The operation of an rf amplifier stage is quite similar in many respects to that of the low-frequency amplifiers that you have already studied. The big difference is in the use of resonant circuits in rf amplifiers. Therefore, before going ahead with our study of the rf amplifier let's learn more about resonant circuits.

For many years we have used the expressions cycles per second, kilocycles per second and megacycles per second to describe the frequency of repetitive waves. For example, a sine wave is a repetitive waveform; it simply repeats itself over and over again. The frequency of the power line voltage, which is a sine wave, is 60 cycles per second.

In place of the expression, cycles per second, a new term, the Hertz, is now being used. Hertz was a physicist who many years ago studied radio wave propagation. No unit in electricity has been named after him and hence the term Hertz was designated as an honor to him and also as a unit of frequency measurement. One Hertz is equal to one cycle per second. 60 Hertz is equal to 60 cycles per second. We usually abbreviate the word Hertz Hz, thus instead of writing the power-line frequency as 60 cps we can write it as 60 Hz.

In addition to the unit Hertz we have the kilohertz and the megahertz. The kilohertz is abbreviated KHz and is equal to 1000 cycles per second or 1 kilocycle per second. The term megahertz is usually abbreviated MHz and is equal to 1,000,000 cycles per second or 1 megacycle per second. Notice that the term Hertz not only identifies the number of cycles, but also the time as one second. Thus to properly describe the power line frequency we can say 60 Hz, but if we use cycles, we must say 60 cycles per second.

You'll find the term cycles per second, kilocycles per second and megacycles per second used in all the older textbooks and magazines. Even some later textbooks still use these units. However, the general trend is toward adopting the new terms Hz, KHz, and MHz. Since you need to be familiar with both sets of units, we will use both in the following lessons. It will be worthwhile to take time now to memorize the equivalents.

1 Hertz (Hz) = 1 cycle per second (1 cps)
1 Kilohertz (KHz) = 1 kilocycle per second (kc ps)
1 Megahertz (MHz) = 1 megacycle per second (1 mc ps)

Practical Facts about Resonant Circuits

In an earlier lesson on resonant circuits you learned that there are two types of resonant circuits, series resonant circuits and parallel resonant circuits. In Fig. 3 we have shown these two resonant circuits. A series resonant circuit is shown in A and a parallel resonant circuit at B.

At this time it might be well to point out again that whether a circuit is a series resonant circuit or a parallel resonant circuit depends not on how the components are connected, but on how the voltage is applied to the circuit. If the voltage is applied in series with the coil and capacitor, as in Fig. 3A, the circuit is a series resonant circuit, but if it is applied across the coil and capacitor in parallel as in Fig. 3B, then the circuit is a parallel resonant circuit. Keep this point in mind; you will run into both series resonant and parallel resonant circuits in all types of communications equipment. It is not always easy to tell whether the circuit is a series resonant or a parallel resonant circuit simply by looking at it, but if you consider how the voltage is applied to the circuit, you can usually tell which type it is without too much difficulty.

Now let us quickly review what we already know about resonant circuits. We know that a circuit will be at resonance when the inductive reactance (X_L) of the coil is exactly cancelled by the capacitive reactance (X_C) of the capacitor. You know that this will occur for any given coil and capacitor at one frequency and only one frequency. The circuit will be resonant at this one frequency. In other words, if we take any coil and connect a capacitor across it, at some frequency that circuit will be resonant because the inductive reactance of the coil will be exactly cancelled out by the capacitive reactance of the capacitor.

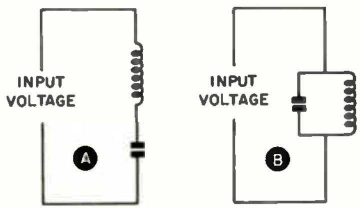


Fig. 3. A series resonant circuit is shown at A; a parallel resonant circuit at B.

The fact that a resonant circuit is resonant at only one frequency does not mean that it will respond only to that exact frequency. In fact, it will respond to a band of frequencies around the resonant frequency. For example, if we apply a voltage to a series resonant circuit and change the frequency of the voltage source, we will find that at resonance we get a maximum current flow through the resonant circuit. If we increase the frequency slightly above the resonant frequency, we will find that the current drops slightly. If we increase the frequency of the voltage source still more, the current will drop a little more. If we increase the frequency still further, the current will drop still further. Similarly, if we reduce the frequency below the resonant frequency, we will find that the current flowing in the circuit is slightly less than the current flowing at resonance. If we reduce the frequency still more, the current will be still less. In other words, at resonance we get a maximum current flow through the resonant circuit and at frequencies either above or below the resonant frequency, the current is somewhat less than it is at the resonant frequency. The further we get away from the resonant frequency, the lower the current will be.

Actually, instead of responding to a single frequency, the resonant circuit will respond to a band of frequencies around the resonant frequency. How wide a band of frequencies it will respond to depends upon the Q of the circuit. If the Q of the resonant circuit is high, any appreciable deviation from the resonant frequency will cause an appreciable change in output. However, if the Q of the resonant circuit is low, there must be a substantial deviation from the resonant frequency before the output of the resonant circuit changes appreciably.

You might at first think that the fact that a resonant circuit will respond to a band of frequencies rather than a single frequency is a disadvantage. However, this is not the case. In fact, if a resonant circuit would pass only a single frequency or a very narrow band of frequencies, then the radio and TV systems that we have today would not be practical. To see why this is so and to help us get a better understanding of what a resonant circuit in an rf amplifier must do, let's once again consider the type of signal that is actually transmitted by a radio or television broadcast station.

SIDEBAND FREQUENCIES

In a radio or television broadcast station, one section of the transmitter generates a radio-frequency signal which is known as the carrier or carrier wave. This is the radio-frequency signal that travels through space and carries either the sound or picture signals being transmitted by the broadcast station. However, a radio-frequency carrier itself is of no value unless we add intelligence to it. The earliest method of using this carrier signal was by interrupting it in a series of dots and dashes to send messages by code. However, even though this is useful in communications work, it is of no value in transmitting radio and television programs for entertainment purposes.

In order to transmit a radio program, the sound or audio signal must be superimposed on the radio frequency carrier. We call this process of superimposing the sound signal on the carrier, modulation.

In the modulation process certain additional frequencies other than the original carrier frequency are produced. For example, let us consider a radio station broadcasting on a carrier frequency of 1000 KHz. If we modulate this signal with a 1000-cycle (1 KHz) signal, we will produce two new frequencies in the modulation process. The 1 KHz audio signal when it is used to modulate the 1000 KHz carrier will produce two new signals, one equal to the sum of the carrier frequency and the audio frequency, and a second signal equal to the difference between the carrier frequency and the audio frequency. In other words, we will produce a signal of 1001 KHz and a signal of 999 KHz. The 1001 KHz signal and the 999 KHz signal are called the sideband signals. The higher of the two sidebands is called the upper sideband and the lower of the two is called the lower sideband.

Now, if we wish to be able to receive this modulated signal on a radio receiver, we must be able to receive not only the original 1000 KHz carrier, but also the two sideband frequencies. The sideband frequencies are the frequencies that actually carry the 1000-cycle audio signal superimposed on the carrier. Therefore, if we had a resonant circuit in our receiver that would respond only to a frequency of 1000 KHz and no other signal frequencies, we would not be able to pick up the 999 KHz and the 1001 KHz signal along with the 1000 KHz signal, and hence we would not be able to receive the modulation on the 1000 KHz carrier.

Fortunately resonant circuits have what is called a bandwidth. This simply means that the resonant circuit will respond to a band of frequencies around the resonant frequency and hence we would have no difficulty designing a resonant circuit that would respond not only to the 1000 KHz carrier signal, but also to the 999 KHz and the 1001 KHz sideband frequencies.

As you will see later, the resonant circuits in radio and television receivers must be able to respond to frequencies substantially above or below the resonant frequency. We will go into this shortly, but first let's look into what we mean by the bandwidth of a resonant circuit and some of the factors that affect the bandwidth.

BANDWIDTH

As we pointed out in the preceding section, a resonant circuit will pass a band of frequencies rather than a single frequency. However, if we move away from the resonant frequency the output that will be obtained from the resonant circuit decreases. In other words, if we are 50 KHz away from the resonant frequency you will not obtain as high an output from the resonant circuit as we would at the resonant frequency.

In Fig. 4 we have shown a voltage response curve of a typical resonant circuit that is resonant at a frequency of 1000 KHz. A response curve is simply a curve that shows how the circuit responds to signals at or near the resonant frequency. Notice that at the resonant frequency of 1000 KHz, the circuit peaks; in other words, the output voltage from the circuit is at its maximum value at a frequency of 1000 KHz. As the frequency is increased or decreased from the resonant frequency, the voltage produced across the circuit begins to go down. Notice, however, that although the circuit is resonant and the voltage is highest at a frequency of 1000 KHz, there is some voltage developed across it when the frequency is as low as 900 KHz and also when it is as high as 1100 KHz. This means that this circuit tuned to 1000 KHz when used in a receiver, will pass, to some extent, frequencies as low as 900 KHz and as high as 1100 KHz. However, the output at these frequencies is substantially below what it is at 1000 KHz.

Engineers have arbitrarily set a standard by which they measure the bandwidth of a resonant circuit. To

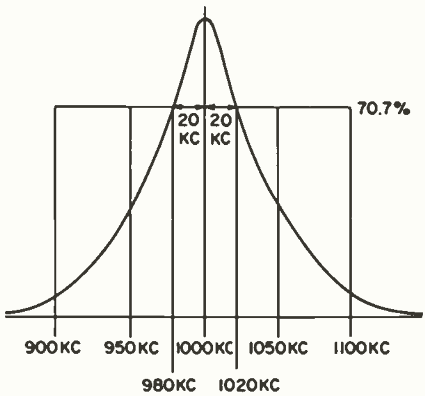


Fig. 4. A response curve showing the two 70.7% points. The bandwidth extends 20 kc on each side of this resonant frequency.

determine the bandwidth of a circuit you find the points at which the output voltage falls to 70.7% of what it is at the resonant frequency. As shown in Fig. 4 these points are at 980 KHz and 1020 KHz. The difference in frequency between these two frequencies is called the bandwidth of the amplifier. In other words, the bandwidth of this amplifier is 40 KHz. It will pass frequencies up to 20 KHz above and down to 20 KHz below the resonant frequency, with at least 70.7% of the output that will be obtained at the resonant frequency.

As in low-frequency amplifiers, the points at which the voltage falls to 70.7% are called the half-power points.

Now the question might come up, do all resonant circuits have the same bandwidth? The answer is no; the bandwidth depends upon the L-C ratio of the resonant circuit and on the Q of the resonant circuit. You will remember that there are many different combinations of L and C that will resonate at a given frequency. Different combinations will have different bandwidths. Furthermore, the Q of the resonant circuit, which is determined chiefly by the inductive reactance of the coil and the resistance in the resonant circuit, will have an effect on the bandwidth of the circuit.

In actual practice the bandwidth of the circuit is sometimes altered by loading the circuit with resistance. In Fig. 5A we have shown a typical parallel resonant circuit. In Fig. 5B we have shown the circuit loaded with a resistor. In Fig. 5C curve 1 shows the bandwidth of the resonant circuit alone and curve 2 shows how the bandwidth is altered by connecting the resistance in parallel with the coil and the capacitor. Notice that curve 2 does not come to as high a peak at the resonant frequency. This indicates that the output from the resonant circuit has been reduced by loading the circuit. However, the bandwidth of the circuit has been increased. Often, it is desirable to sacrifice output in order to obtain a wide bandwidth; this is frequently accomplished by loading the circuit with a resistance. The lower the value of resistance, the lower the output and the wider the bandwidth of the resonant circuit.

Loading resonant circuits by connecting resistors across them reduces the Q of the circuit. It is seldom necessary to do this in radio receivers because the required bandwidth can be obtained by the correct design of the tuned circuit. In fact you do not want too wide a bandwidth because you must have sufficient selectivity to be able to select the one signal you want from among many signals picked up by the receiver. The ideal arrangement is a circuit with a wide enough bandwidth to pass all the sidebands being transmitted by the station and enough selectivity to reject all signals beyond the sideband frequencies. However, you will find that in TV receivers the resonant circuits are frequently loaded by connecting resistors across them in order to get the wide bandwidth needed to pass all the sidebands that carry the picture and color detail.

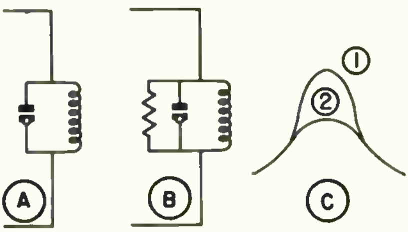


Fig. 5. The response curves at C show the effect of loading a parallel resonant circuit as at B.

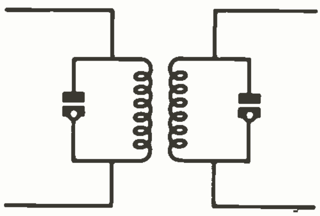
COUPLING RESONANT CIRCUITS

RF stages are frequently coupled together by means of rf transformers. A typical example of this type of transformer is an i-f transformer. An i-f transformer consists of two resonant circuits as shown in Fig. 6A. One resonant circuit is used as the load in the output circuit of one stage and the other resonant circuit is used as the input in the following stage. This double resonant circuit helps to improve the selectivity of the receiver. Selectivity is the ability of the receiver to receive the desired signal and reject undesired signals. As we pointed out, the receiver must have enough selectivity to reject undesired signals from stations operating near the frequency of the desired signal.

A photograph showing construction of an i-f transformer is shown in Fig. 6B. Notice that the two coils are wound on a round cardboard form. The coils are not wound on top of each other nor are they placed exactly side by side; you can see that there is some spacing between the two coils. However, in spite of this spacing, the coils are close enough so that they are inductively coupled together.

The exact spacing between the two coils affects the degree of coupling between the two coils. In other words, if the spacing is great, not all the lines of force produced by the primary will cut through the secondary. There is some maximum spacing beyond which some of the lines of force produced by the primary will be lost and will not cut the secondary.

An example of the effect of varying the spacing between the two coils is shown in Fig. 7. If we apply a signal from a variable frequency generator to the primary and measure the output voltage across the secondary as the frequency is varied, we would obtain data from which we could plot curves like those shown. Curve 1 shown in Fig. 7A represents a certain spacing between the two coils, where the coils are pushed quite far apart. As the coils are pushed closer together the output across the secondary will increase until eventually a point is reached where maximum output is obtained as shown in curve 2 at B. If the coils are pushed still closer together we find that the output at the resonant frequency drops somewhat, as shown in curve 3 in Fig. 7C, but rises slightly above and slightly below the resonant frequency, producing two humps in the response curve with a valley in between them. This is often called a double-hump curve. If we push the coils still closer together, the response curve will be still broader. The two peaks have a tendency to move somewhat farther apart and drop in height, and the valley in the center becomes more pronounced.

A


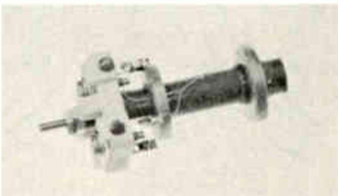
B


Fig. 6. A schematic diagram of an i-f transformer is shown at A; a photo at B.

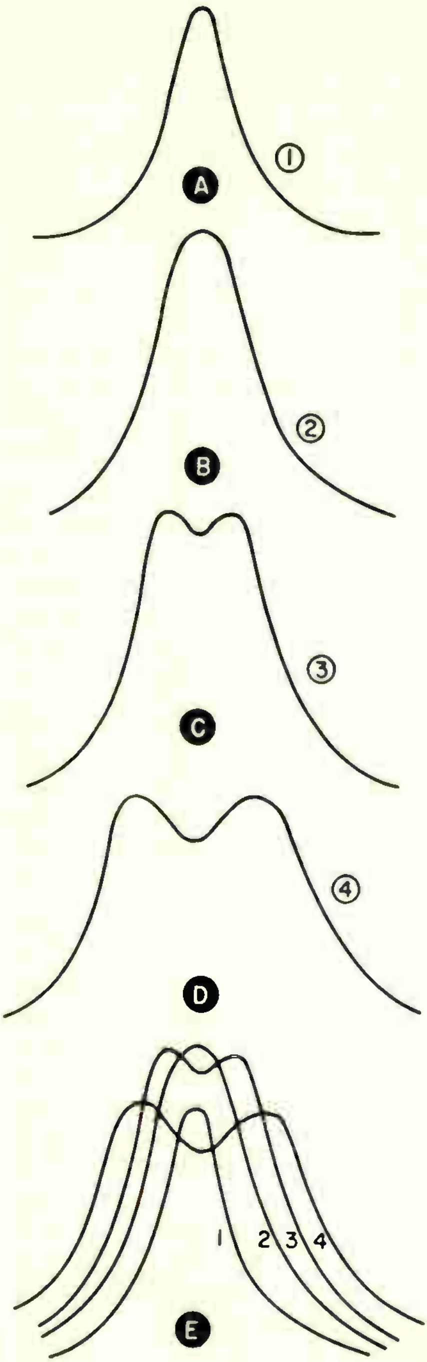


Fig. 7. Response curves showing how coupling affects bandwidth and response.

The four curves shown in A, B, C, and D have been superimposed on a single drawing at E so you can see the effect of changing the spacing between the coils.

As the coils are pushed closer together some point is reached where the output from the secondary is at a maximum value. Reducing the spacing beyond this point results in the two humps in the curve appearing. At the point where the output is maximum and just before the humps in the curve begin to appear, we have what is called critical coupling. If the coils are spaced farther apart than this particular spacing we say that they are under-coupled, in other words the coupling is less than critical coupling. If the coils are pushed closer together than this spacing, we say that the circuits are over-coupled, in other words the coupling is tighter or closer than critical coupling.

As a technician you will not have to adjust the spacing between the primary and secondary coils on an i-f transformer. They will already be adjusted for you by the manufacturer. The i-f transformers in a broadcast-band radio receiver are usually adjusted at, or slightly beyond, the critical coupling point. The i-f transformers in a communications receiver are usually adjusted below the critical coupling point. The i-f transformers in an FM receiver or in the FM sound section of a television receiver are adjusted somewhat beyond the critical coupling point.

In the early days of television, double-tuned i-f transformers were used in the video (picture) i-f amplifier stages. These transformers were very heavily overcoupled, and in addition, they were usually loaded by resistors in order to reduce the Q of the primary and secondary circuits and obtain a very wide bandwidth.

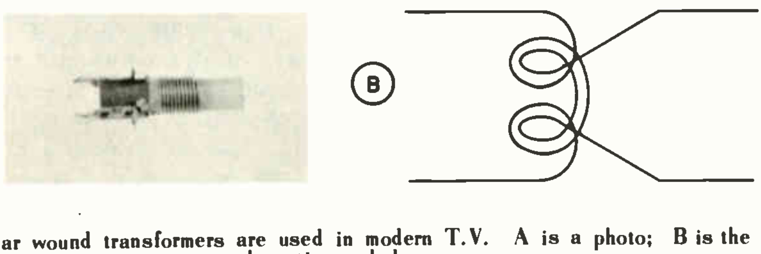
However, modern TV receivers usually employ what is called bifilar wound transformers. A transformer of this type is shown in Fig. 8A. Notice that the secondary winding is wound directly over top of the primary winding in order to obtain as tight coupling as possible between the two windings. Bifilar windings are usually represented schematically by the symbol shown in Fig. 8B. The fact that the coils are interwound so that the primary turns are mixed directly with the secondary turns is schematically represented. The advantage of the bifilar wound transformer is that the coupling between the two coils is very tight. A single slug is used to adjust the resonant frequency of the circuit. The two coils are so tightly coupled that they act like one coil. The output capacity of the stage driving the primary circuit is in effect directly in parallel with the input of the second stage. The net result is that the single slug can be used to adjust the two circuits to resonance at the same frequency. Bifilar wound coils are excellent where a wide frequency band is needed. This situation is encountered in the video i-f amplifier of a television receiver as you will see later.

There are reasons for using the various types of coupling in each of the particular applications mentioned. In a radio receiver designed to receive radio signals on the standard broadcast band the receiver must be able to pass the carrier wave of the broadcast station plus its sidebands. In the standard broadcast band the sidebands do not extend too far above and below the carrier frequency and therefore an extremely wide bandpass is not required.

In communications receivers the primary purpose of the receiver is to be able to receive information. Often the station that you are listening to may be operating very close to other more powerful stations. Here you want as much selectivity as you can get. Therefore the i-f coils are set either at or below critical coupling in order to get as much selectivity as possible from the transformer. If some of the sidebands of the signal are missing this will not be too important. In communications circuits where communications receivers are used, you are usually interested in receiving voice transmission, and the frequency range of the human voice is comparatively limited.

A much wider bandwidth is used in FM transmitters than in AM transmitters. In FM transmissions, the amplitude of the signal transmitted by the FM station does not vary. Instead, the frequency of the signal is varied. It can be varied in the standard FM broadcast station as much as 75 kilohertz above or 75

A



B

Fig. 8. Bifilar wound transformers are used in modern T.V. A is a photo; B is the schematic symbol.

kilohertz below the center frequency or resting frequency as the carrier frequency of the FM station is called. The rate at which the signal varies above and below the resting frequency is determined by the frequency of the audio signal being transmitted. How far above and below the resting frequency the signal varies is determined by the amplitude of the audio signal. A loud signal can produce sidebands two hundred KHz wide. Therefore the FM broadcast receiver must be capable of passing all these sideband frequencies, particularly if the FM station is a high-fidelity station.

The FM sound system in a television receiver is called narrow band FM. Here the maximum deviation above and below the sound carrier frequency is limited to 25 kilohertz. The rate at which the carrier resting signal is varied above and below the carrier frequency is again determined by the frequency of the signal being transmitted and the amount varied with the amplitude. Even with the deviation limited to 25 KHz, sideband frequencies considerably in excess of 25 KHz are easily produced. The sound i-f transformer must pass this band of frequencies in a TV receiver or the sound signals will be distorted.

In a TV receiver the picture signal consists of a carrier plus the modulation information on it. The modulation signal may produce sidebands up to 4 megahertz wide. Therefore in order to reproduce the sidebands the i-f bandwidth on the TV receiver must be comparatively wide. In color TV, a subcarrier having a frequency of about 3.5 MHz is used. This produces a video i-f signal that is 3.58 megahertz lower in frequency than the video i-f carrier. If the i-f bandwidth of the receiver is not wide enough to pass the video carrier signal, sidebands and the color subcarrier, then the set could not reproduce a color picture. Later in your course when you study television in detail you will understand why it is important that the video i-f amplifier of a television receiver have such a wide bandwidth.

SERVICING NOTES

We have already pointed out that a high Q resonant circuit not only has better selectivity than a low Q circuit, but also for a given input has a higher output. Thus if a piece of electronic equipment is designed with high Q resonant circuits, anything that lowers the Q of the circuits will reduce the gain and the selectivity of the equipment. The chances are that you may not notice a change in selectivity, but it is quite likely that you will notice that the gain of the equipment has fallen off appreciably.

There are several things that might reduce the Q of a resonant circuit. Coils can absorb moisture. Often in damp weather dust or other particles will settle on the coil and adhere to it. Both of these effects will introduce resistance in the circuit, lowering the Q of the circuit and changing both the output and the selectivity.

Sometimes the low Q is due to a poorly soldered connection in a resonant circuit. If the connection is not properly soldered, it may work well for a while and then develop trouble. Poor connections of this type often get through the factory inspection. Holding a hot soldering iron on suspected connections until the solder flows freely over the leads will usually clear up this type of trouble.

You will frequently find in servicing radio receivers that have been in use for many years, that even though you replace defective tubes so that you have all good tubes in the set and all operating voltages throughout the set are normal, the gain of the receiver is not all it should be. When realigning the set does not clear up the trouble, the difficulty is often due to the fact that the Q of one or more of the resonant circuits has fallen off because of moisture or dirt absorption. In a situation like this you must find the resonant circuit causing the trouble and replace it to restore the equipment to its original gain and selectivity.

Trouble of this type is not often found in transistor radios or in TV receivers. As we pointed out, this is something that occurs after a receiver has been used for many years. Most transistor radios are too new to have developed this kind of trouble. The trouble is not found in TV receivers, because usually the resonant circuits in the television receiver are very heavily loaded in order to give the wide bandwidth needed to pass the picture signal.

Resonant circuits are tuned to resonance by adjusting the capacity or the inductance in the circuit so that the inductive reactance of the coil is exactly equal to and cancels out the capacitive reactance of the capacitor at the resonant frequency. Any further change of either the inductance or the capacitance in the circuit will shift the resonant frequency of the circuit.

Both tubes and transistors have a certain amount of internal capacitance. For example, in an i-f tube there is capacity between the grid and the cathode and between the grid and the screen. The screen is grounded insofar as rf is concerned so that the grid-to-screen capacity is effectively placed between the grid and ground. The cathode of the tube is usually operated at rf ground potential so that the grid-to-cathode capacity is placed directly between grid and ground. Thus if the input of the tube is connected across a resonant circuit, the tube capacities affect the resonant circuit. If an i-f amplifier is aligned with one tube in a circuit and then a different tube is installed, it is likely that the new tube will not have exactly the same input capacity as the old one. As a result, installing the new tube in the circuit will slightly detune the resonant circuit. Radio receivers designed for broadcast band reception are usually fairly broad, so the change in capacity will not be enough to cause trouble. However, if you change two or three tubes in an i-f amplifier of a TV receiver the resonant frequency of the various circuits in the i-f amplifier may be altered enough to appreciably alter the bandwidth of the i-f amplifier so that it can no longer pass all the sideband frequencies and there will be some loss in picture detail. In transistor equipment you have capacity between the emitter and the base and between the base and collector which can have the same effect on resonant circuits.

In communications receivers designed to provide good selectivity in order to separate stations operating very close together, changing one or more tubes or transistors in one of the tuned circuits such as the rf amplifier, mixer, or i-f stages will usually alter the resonant frequency of the particular circuit involved to such an extent that the selectivity of the receiver will suffer. This is particularly true of stages designed to operate at high frequencies.

Disturbing the wiring in a resonant circuit designed for operation at a high frequency may change both the inductance and capacity in the circuit. Even a straight short piece of wire has a certain amount of inductance. At the comparatively low frequencies used for standard radio broadcasting a straight short piece of wire has so little inductance that it can be ignored. However when you get up into the higher frequencies such as those used for TV and FM broadcasting, even the smallest inductance becomes important. Changing the length of a wire in a critical circuit may have some effect on the resonant frequency because the inductance in the circuit is changed. Also moving the position of a wire in a resonant circuit operating at a high frequency and pushing it closer to a metal chassis or moving it further away from the metal chassis may change the capacity of the circuit. There is always capacity between the wires in a resonant circuit and the chassis or ground. Moving the wires around in a broadcast-band receiver is not likely to cause any trouble, but changing the position of a wire in a high-frequency circuit will frequently have an appreciable effect on the circuit.

These points are very important and are worth remembering when working on resonant circuits found in TV receivers or any other equipment designed for high-frequency operation. When it is necessary to make repairs on components in or near resonant circuits, it is a good idea to avoid disturbing the parts and leads as much as possible. If you have to move a lead to get at another part in order to replace it, try to put the lead back in as close as possible to the original position. Often this is an unnecessary precaution, but it is a precaution that can keep you out of difficulty in some critical circuits.

Much of the material covered in this section has been a review for you, but now that you have the important points about resonant circuits fresh in your mind you are ready to go ahead with your study of rf amplifiers.

SELF TEST QUESTIONS

(a) What is the basic difference between the circuit used in an rf amplifier and the circuit used in an audio amplifier?

(b) What is the basic difference between a series-resonant and a parallel-resonant circuit?

(c) What are sideband frequencies?

(d) What do we mean by the bandwidth of a resonant circuit?

(e) On what factors does the bandwidth of a resonant circuit depend?

(f) How can the bandwidth of a given parallel-resonant circuit be increased?

(g) What do we mean by critical coupling between two coils?

(h) What do we mean by overcoupling?

(i) What happens to a resonant circuit that absorbs moisture, and dust and particles settle on the coil and adhere to it?

(j) Why should you avoid moving any of the wires in the rf section of a TV receiver?

Resonant circuits are extremely important in rf amplifiers. If you think you may have forgotten some of the details you learned about resonant circuits it would be a good idea to spend some time reviewing to be sure you understand and remember the important characteristics of both series and parallel-resonant circuits. This will help you in these sections of the lesson to follow where we'll go into detail about both vacuum tube and transistor radio-frequency amplifiers.

Radio-Frequency Voltage Amplifiers

A block diagram of a superheterodyne receiver is shown in Fig. 9. You have seen this block diagram before. The rf amplifier is the stage connected to the antenna. The weak signal from the antenna is fed to the rf stage where it is amplified and then fed to the mixer. The rf stage also has tuned circuits so that it provides a certain amount of selectivity. In other words, it helps select the desired signal and reject undesired signals.

The signal from the rf amplifier is fed to a mixer-oscillator stage. In this stage the incoming signal is mixed with a locally generated signal and a new signal is produced. This signal that we are interested in is called the intermediate-frequency signal. It is equal in frequency to the difference between the incoming signal and the oscillator signal. This intermediate-frequency signal, or i-f signal as it is called, is fed to an i-f amplifier where it receives further amplification. The signal is still a radio-frequency signal so strictly speaking the i-f amplifier is a radio-frequency amplifier. However, it operates at much lower frequencies than the rf amplifier that precedes the mixer. From the i-f amplifier the signal is then fed to a second detector where the intelligence on the carrier wave is removed and then it is fed to a low-frequency amplifier. In the case of a radio receiver the signal is then fed to a speaker; in the case of a television receiver the sound signal is fed to a speaker and the picture signal is fed to the picture tube and to various other circuits you will study later.

The circuit we are going to be concerned about in this part of this lesson is the rf amplifier used between the antenna and the mixer. This stage is omitted in some of the small low-cost radio receivers, but it is invariably found in the better receivers and television receivers.

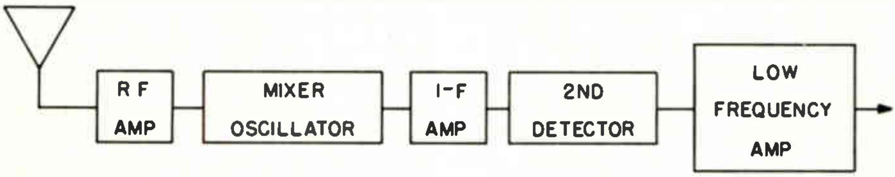


Fig. 9. A block diagram of a typical superheterodyne receiver.

The stage is a voltage-amplifier stage. In later sections of the lesson you will study i-f amplifiers and in another section rf power-amplifier stages.

The rf voltage amplifiers used in radio and television receivers are all Class A amplifiers. Some use pentode tubes, some use triode tubes, and some use transistors. You will study all three types in this section of the lesson.

PENTODE RF AMPLIFIERS

A typical rf amplifier using a pentode tube is shown in Fig. 10. This is the type of amplifier that is used between the antenna and the mixer in a communications-type superheterodyne receiver, in a few of the better broadcast-band receivers, and in many FM receivers.

In the circuit shown in Fig. 10, the signal picked up by the antenna causes a current to flow through L_1 which is the primary of the antenna transformer T_1. L_1 and L_2 are wound on the same form and are inductively coupled together. Thus the magnetic field produced by the current flowing through L_1 cuts the turns of L_2 and induces a voltage in series with it. L_2 is tuned to resonance by the capacitor C_1. Since the voltage induced in L_2 is induced in series with it, the combination of

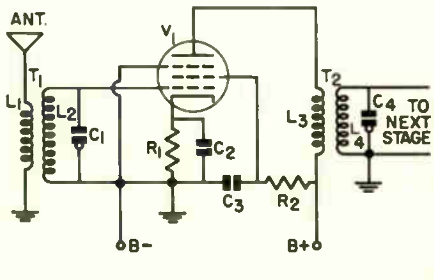


Fig. 10. A typical rf amplifier using a pentode tube.

L_2 and C_1 forms a series-resonant circuit. The frequency to which this circuit is tuned can be altered by changing the capacity of C_1. This is done by rotating the dial on the receiver, which causes the rotor or moving plates of the tuning capacitor to move in and out between the stator (stationary) plates of the capacitor.

In some receivers instead of using a variable capacitor such as shown in Fig. 10, a trimmer capacitor is used and a powdered iron slug is used that can be moved in and out of L_2. A trimmer capacitor is a capacitor that is adjusted by means of a screwdriver to adjust the stage at the high-frequency end of the band to be covered. The actual change in the resonant frequency of the tuned circuit required to tune the receiver across the band is then accomplished by moving the slug in and out of the coil L_2. This type of tuning, which is often called permeability tuning, is found more frequently in FM receivers than in broadcast and communications-type receivers.

It makes comparatively little difference whether the frequency of the tuned circuit is varied by varying the inductance of the coil or capacity of the capacitor. Either will change the resonant frequency so that the rf stage can be tuned to the frequency of the signal to be amplified.

We mentioned that L_2 and C_1 form a series-resonant circuit. We have mentioned before that when the voltage is induced in the secondary winding of a transformer, in a circuit of this type, that the circuit is a series-resonant circuit. Fig. 11 is an example of what happens. A small voltage is induced in each turn of the coil. These voltages add up to give you the total voltage induced in L_2. It is more or less like a series of

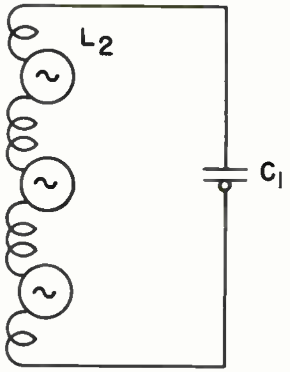


Fig. 11. Voltage induced in L_2 is induced in series with the turns of L_2.

small generators placed in the coil in series with the various turns as shown in Fig. 11. It is important to realize that this type of circuit is always a series-resonant circuit and never a parallel-resonant circuit; it will help you understand better what is happening in the circuit.

You will remember that one of the characteristics of a series-resonant circuit is that there is a high circulating current in the circuit and that there will be a resonant voltage step-up across the coil and across the capacitor. Thus we have the weak signal voltage picked up by the antenna being stepped-up by the resonant circuit. The voltage across the coil and across the capacitor in the resonant circuit is then applied between the grid and cathode of V_1. It is applied directly to the grid of the tube and to the cathode through the cathode-bypass capacitor C_2.

The radio-frequency signal applied between the grid and cathode of V_1 will cause the plate current flowing through V_1 to vary at the same rate as the incoming signal. Remember that L_2 and C_1 will resonate at only one frequency and signals at this particular frequency will normally be stepped up and be much stronger than any other signal.

The varying current flowing through the tube will flow through L_3, which is the primary winding of T_2. L_3 is inductively coupled to L_4 and hence a voltage will be induced in L_4. Again L_4 and C_4 make up a series-resonant circuit. The combination of L_4 and C_4 will be tuned to the same frequency as L_2 and C_1 and hence this resonant circuit will give the signal a still further build up, and at the same time help to reject any signals of a frequency other than the resonant frequency that happened to get by L_2 and C_1 to the grid of V_1.

Thus we have three things increasing the signal voltage. We have the resonant voltage step-up in the L_2 C_1 series-resonant combination, we have the voltage gain that can be obtained from the tube V_1 and a still further voltage step-up in the resonant circuit consisting of L_4 and C_4.

Notice that screen grid of the pentode tube is connected to ground through the capacitor C_3. C_3 is chosen so that at radio frequencies it has a very low reactance, or in other words it acts like a short circuit insofar as radio frequencies are concerned. Thus the screen is said to be at rf ground potential. This simply means that as far as the rf signal is concerned, the screen grid might just as well be connected directly to ground.

The screen grid shields the grid of the tube from the plate so that little or no energy can be fed from the plate of the tube back to the grid of the tube. It is important that we avoid feeding any energy from the plate of the tube back to the grid, otherwise we may have trouble with the stage going into oscillation.

We mentioned that this type of rf amplifier may be found in broadcast band receivers, communication receivers and FM receivers. The main difference between the rf amplifiers found in these different pieces of equipment will be in the number of turns on the coils. In an FM receiver that operates in the vicinity of 100 MHz, the coils L_1, L_2, L_3 and L_4 will usually consist of only one or two turns of a rather large diameter wire. In the broadcast band, the coils will consist of 100 or more turns of a comparatively fine wire. In communications receivers you will find coils all the way between the coils with 100 or more turns used in broadcast receivers to the coils with even fewer turns than those found in FM receivers. The exact number of turns on the coil will depend upon the frequency band the receiver is to cover.

Another pentode rf amplifier is shown in Fig. 12. This is identical to the amplifier shown in Fig. 10 except the capacitor C_5 and resistor R_3 have been added. C_5 is a comparatively large capacitor so that at the signal frequency its reactance is almost zero. Therefore as far as L_2 is concerned it is connected directly to B- and to the rotor of C_1. The signal sees no opposition because the reactance of C_5 is so low. R_3 is a comparatively large resistor so there is no signal fed from R_3 into the circuit we have marked avc.

The purpose of this circuit is to enable us to apply a variable negative voltage to the grid of V_1. The negative voltage applied to the grid of V_1 will control the gain of the tube. The higher the negative voltage, the lower the gain of the tube. The voltage is usually obtained from the detector circuit and the strength of this voltage will depend upon the strength of the signal. This voltage is referred to as the automatic volume control voltage and is used to vary the gain of the stage. If the input signal is very strong, the avc voltage developed will be high and will tend to reduce the gain of the stage. On the other hand, if the signal received is weak, very little avc voltage will be developed and the stage will operate at close to its maximum gain. You will see in a later lesson how this avc or automatic volume control voltage is developed.

As a point of interest, avc is the term used in radio receivers. Actu-

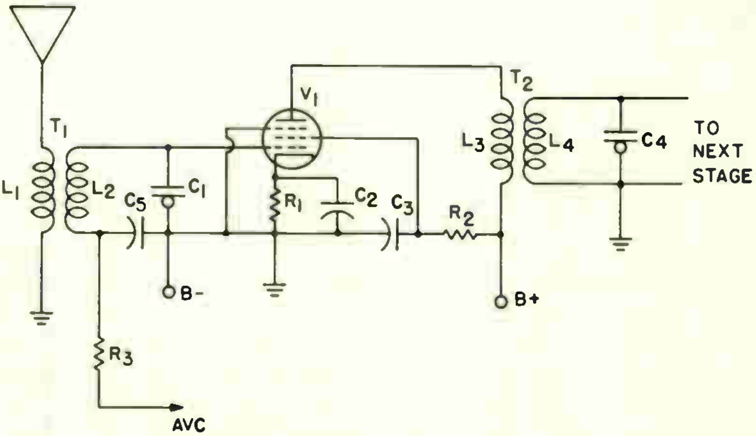


Fig. 12. A pentode rf amplifier modified for automatic gain control.

ally, this is not an automatic volume control, but rather it is an automatic gain control. In television, we use the term automatic gain control, abbreviated agc rather than avc. In television agc is much more suitable because varying the gain of the rf amplifier affects both the picture and sound signals.

TRIODE RF AMPLIFIERS

One of the disadvantages of pentode rf amplifiers is that they generate considerable noise within the tube itself. Part of this noise is often caused by changes in the division of the cathode current between the plate and screen of the tube. Small changes occur at random. These changes are too small to be detected in most cases, but they result in noise being generated within the tube. In the case of radio stations operating in the broadcast band and FM stations, usually the signals are so strong that they simply override this noise and it does not cause any trouble. However, in TV receivers the noise may become objectionable and may appear in the picture, particularly on weak signals. This problem can be overcome by using a triode rf amplifier.

Triode tubes have been especially designed for use as rf amplifiers in vhf tuners of television receivers. The vhf tuner in a television receiver covers channels 2 to 13. A typical triode rf amplifier such as might be found in a television receiver is shown in Fig. 13.

In TV receivers the lead-in or wire used to connect the antenna to the receiver is usually what is called a balanced wire or cable. This cable has two conductors and neither is at ground potential. Each is operated above ground. The resistance or impedance from either wire to ground is the same. We call this type of line a balanced transmission line.

The balanced transmission line is connected to the antenna terminals of the receiver. In the circuit shown in Fig. 13 it is fed to T_1. T_1 is what is called a balun. The purpose of the balun is to match the balanced transmission line to the unbalanced input circuit. An unbalanced input circuit is a circuit in which one side of the circuit is grounded. T_1 serves the

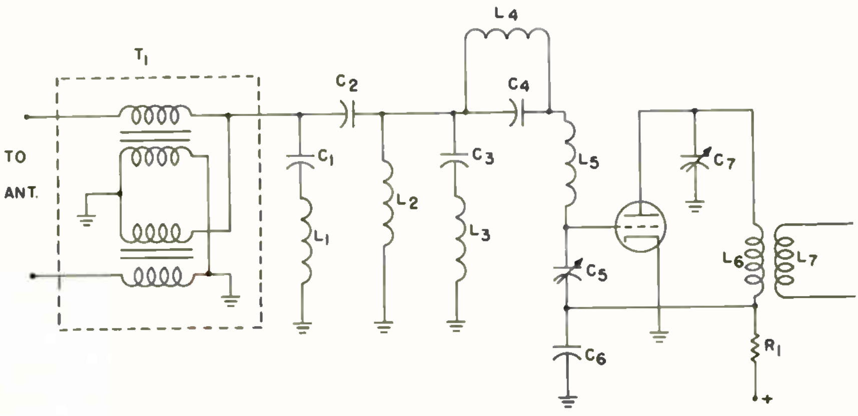


Fig. 13. A triode rf amplifier.

purpose of matching the balanced line to the unbalanced input of the receiver.

The capacitors C_1, C_2 and C_3 along with coils L_1, L_2 and L_3 form what is called a high-pass filter. A high-pass filter is a filter that passes signals above a certain frequency and rejects signals below this frequency. The lowest vhf TV channel is channel 2. The frequency of this TV channel is 54 to 60 megahertz. The high-pass filter is designed to cut out all signals below 54 megahertz and pass signals above 54 megahertz. Thus strong signals from nearby broadcast stations operating the standard broadcast band are prevented from reaching the rf amplifier and causing interference.

The combination of L_4 and C_4 form a parallel-resonant circuit. This circuit is usually a uhf trap and is designed to prevent interference from uhf television signals from reaching the grid of the rf circuit.

L_5 and C_5 form a series-resonant circuit that is resonant to the frequency of the channel received. C_6 is a very large capacitor and insofar as the resonant circuit is concerned has no effect on the circuit.

The signal applied to the grid of the tube is amplified and fed to L_6. The combination of L_6 and C_7 forms a parallel-resonant circuit which is resonant to the frequency of the TV channel.

B+ is applied to the plate of the tube through R_1. The lower end of L_6 is not at ground potential. It is bypassed through C_6. However, C_6 does not bypass all of the signal, but part of it is fed back through C_5 to the grid of the tube. The purpose of feeding this signal back to the grid of the tube is to make up for any signal fed from the plate of the tube back to the grid of the tube through the tube itself. The signal fed from the plate of the tube back to the grid through the tube could cause the stage to go into oscillation. The signal fed from the lower end of L_6 through C_5 to the grid of the tube is 180° out-of-phase with the signal fed through the tube and cancels this signal thereby preventing oscillation. You will remember that when you studied triode tubes before, we pointed out that they would oscillate in rf circuits unless they were neutralized. This is a neutralizing circuit.

In the early days of radio, triode amplifiers were used in broadcast-band receivers, but since the pentode tube was invented, triode tubes have not been used for this purpose. However, you will find triodes in the rf amplifier of TV receivers where the low-noise characteristic of the triode is an advantage over a pentode. The circuit in Fig. 13 is typical. You do not have to memorize this circuit. We will go into it in more detail later when you study TV, but at least you should have an idea of the general circuit configuration and notice that except for the balun and trap in the input circuit, the circuit is not too different from the pentode rf amplifier circuit. Of course, as far as the triode amplifier itself is concerned, the balun, the high-pass filter and the uhf trap could be eliminated and a signal fed directly to the resonant circuit and to the grid of the tube. These other components are needed because of the circumstances under which the amplifier is used.

TRANSISTOR RF AMPLIFIERS

A typical transistor rf amplifier such as might be found in a radio receiver is shown in Fig. 14. The circuits use an NPN transistor in a common-emitter circuit.

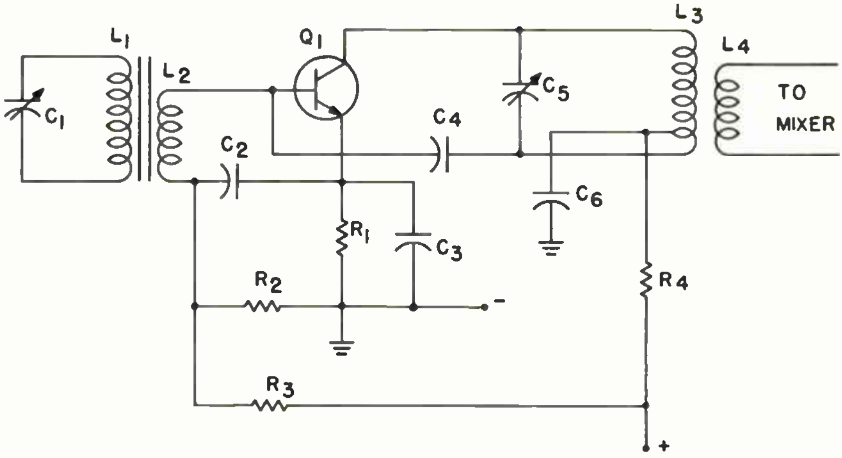


Fig. 14. A transistor rf amplifier.

L_1 is a coil wound on a ferrite rod. A ferrite rod is a rod made of a powdered iron-type material held together by a suitable binder. The coil is wound on this rod to get a very high Q. It is tuned to resonance by C_1 so that the combination of L_1 and C_1 form a series-resonant circuit. The circuit can be tuned to the frequency of the desired station. L_2 is inductively coupled to L_1. Since L_2 has fewer turns than L_1, there will be a current step-up, in other words the current flowing in L_2 will be higher than the current flowing in L_1. This serves two useful purposes; it provides a higher current for the emitter-base circuit of the transistor Q_1, and prevents the transistor from loading the resonant circuit. If it did load the resonant circuit excessively it would lower the Q of the coil and in so doing reduce the selectivity of the circuit.

The signal induced in L_2 is applied to the base of the transistor and to the emitter through C_2. C_2 is selected to have a low reactance at the signal frequency.

The forward bias for the emitter-base junction is provided by the voltage divider consisting of R_2 and R_3. R_3 is much larger than R_2 so that the base is made only slightly positive with respect to the emitter. R_1 is placed in the emitter circuit between the emitter and ground in order to stabilize the bias with temperature increases.

The signal applied between the emitter and base of the transistor varies the forward bias across the emitter-base junction which causes the number of electrons crossing the junction and reaching the collector to vary. The varying current flowing through the transistor reaches the parallel-resonant circuit consisting of L_3 and C_5. The current flows through L_3 to the tap on the coil and then through R_4 back to the positive terminal of the power supply.

Notice that the coil L_3 is tapped. The tap on L_3 is essentially at signal-ground potential. A voltage is induced in the lower end of L_3 which will be 180° out-of-phase with the voltage at the upper end of L_3. The voltage from the lower end of L_3 is fed through C_4 back to the base of the transistor. This voltage is fed back to neutralize or cancel any voltage that is fed from the collector of the transistor to the base through the collector-base capacity. This will prevent the stage from oscillating.

L_4 is inductively coupled to L_3, and since it has fewer turns than L_3 we have in effect a step-down transformer. Thus the input of the mixer transistor does not load the resonant circuit in the collector circuit of the rf amplifier.

In many ways the transistor rf amplifier is similar to the triode tube rf amplifier. Both may go into oscillation unless steps are taken to neutralize the feedback voltage through the device.

Transistors are also used as vhf rf amplifiers in TV receivers. A typical transistor vhf amplifier for a television receiver is shown in Fig. 15. Here you can perhaps see more closely the similarity between the transistor circuit and the triode tube circuit.

In Fig. 15, T_1 is a balun and its purpose once again is to match the balanced transmission line from the antenna to the unbalanced input of the rf amplifier.

In this circuit the combination of L_1 and C_1 form a parallel-resonant circuit. This circuit is resonant at approximately 41 MHz. The sound i-f frequency in the TV receiver is 41.25 MHz. The purpose of this circuit is to prevent any interference by a nearby station operating on 41 MHz from getting to the rf stage and on through into the sound circuits of the receiver.

The combination of C_2 and L_2 form a series-resonant circuit at 45.75 MHz. The picture i-f frequency in the TV receiver is 45.75 MHz. A series-resonant circuit offers a low resistance to signal frequencies near 45.75 MHz. Therefore this circuit will prevent interference from a nearby station operating on or near the video i-f from getting into the picture and causing interference.

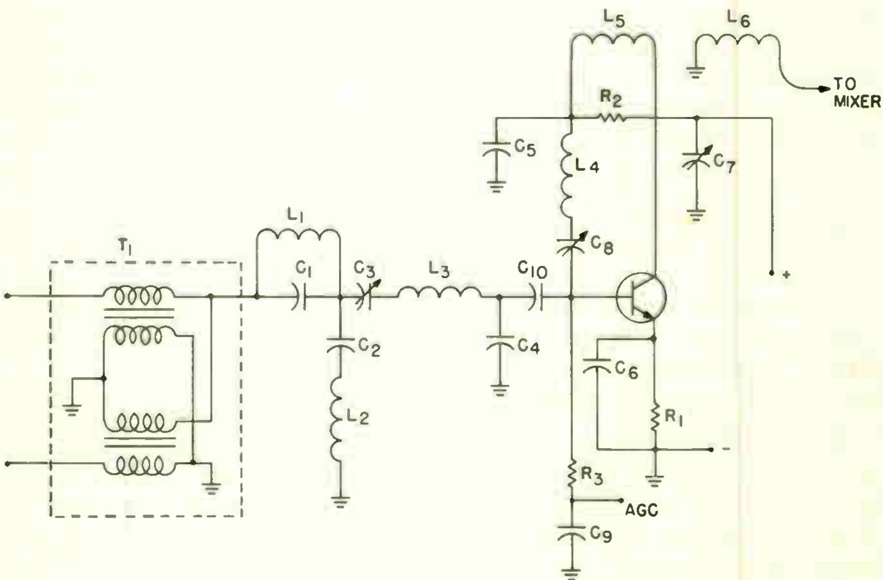


Fig. 15. Transistor vhf amplifier.

C_3 and L_3 form a series-resonant circuit which is tuned to the frequency of the desired TV station. The signal is fed through the capacitor C_{10} to the base of the transistor. The emitter of the transistor is at signal-ground potential, the resistor R_1 in the emitter circuit is bypassed by capacitor C_6.

The signal voltage fed between the base and emitter of the transistor causes the electron current through the transistor to vary and this causes the current flowing from the collector of the transistor and through L_5 to vary. The current from the collector flows through L_5, through R_2 and on to the + terminal of the power supply.

L_5 is inductively coupled to L_6; therefore a voltage will be induced in L_6. This voltage is fed to the mixer stage which follows the rf stage.

The signal at the lower end of L_5, that is the opposite end from the collector, is bypassed to ground through C_5. However, C_5 is not a perfect bypass and part of the signal is fed through L_4 and C_8 back to the base of the transistor. This is the neutralizing voltage which makes up for and cancels the signal fed across the collector-base capacity of the transistor back into the base circuit. Without this neutralizing circuit the rf stage would oscillate.

The forward bias for the transistor is applied through the resistor R_3 to the base. This bias circuit connects back to the automatic gain control circuit in the television receiver. The automatic gain control circuit regulates the gain of the rf stage automatically. In the case of a strong signal, the automatic gain control voltage reduces the gain of the rf stage, and in the case of a weak signal, it allows the stage to operate at maximum gain. You will go into automatic gain control circuits in detail later. Using an automatic gain control circuit to control bias across the emitter-base junction of the transistor will provide much more satisfactory results than applying a fixed forward bias across the junction.

FIELD-EFFECT TRANSISTOR AMPLIFIERS

As you learned earlier, the field-effect transistor combines many of the desirable characteristics of the vacuum tube, along with those of the transistor. Therefore it is reasonable that the field-effect transistor would make an excellent rf amplifier. Both the junction type and the insulated-gate type can be used as rf amplifiers.

Fig. 16 is a diagram showing a junction-type N channel field-effect transistor used as an rf amplifier.

In the circuit shown in Fig. 16, the source is connected to the negative side of the power supply and the drain to the positive side. A negative bias is applied to the gate through L_2 and R_1 from the agc terminal. The negative bias on the gate sets the current flow through the channel from the source to the drain of the transistor.

When a signal is received by the

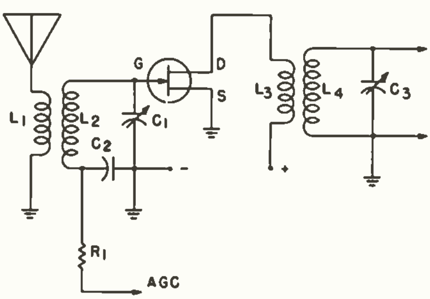


Fig. 16. An rf amplifier using a junction field-effect transistor.

antenna it causes a current to flow through L_1. L_1 is inductively coupled to L_2 and the voltage is induced in series with the turns of L_2. L_2 and C_1 form a series-resonant circuit which is tuned to the frequency of the incoming signal.

The induced voltage across L_2 is applied between the gate and ground in series with the negative bias applied to the gate. This voltage will add to or subtract from the gate bias depending upon the polarity of the signal voltage. The varying rf voltage will modulate the current flowing from the source to the drain, by varying the effective width and hence the resistance of the channel. Thus we have the signal voltage applied to the gate of the transistor causing substantial variations in the current flowing through the transistor.

The varying signal current from the drain of the transistor flows through L_3. L_3 is inductively coupled to L_4 and hence a signal voltage is induced in L_4. L_4 and C_3 form a series-resonant circuit which is tuned to resonance at the same frequency as L_2 and C_1. The output from the resonant circuit consisting of L_4 and C_3 can then be fed to another rf amplifier, to a mixer or to a detector.

The high input resistance of the junction-type field-effect transistor makes this transistor ideally suited for use as an rf amplifier. In operation, it actually resembles very closely a pentode tube. While the transistor we have shown in Fig. 16 is an N-channel transistor, a P-channel transistor can be used just as well. In this case the polarity of the voltages would be reversed and also a positive voltage would be applied to the gate in order to reverse bias the junction and prevent any current flow in the gate circuit.

There will be some reverse current flow across the junction in a junction-type field-effect transistor. This will have the effect of lowering somewhat the input resistance of the transistor. This can be overcome by the use of an insulated-gate field-effect transistor in a circuit such as shown in Fig. 17. The circuit here shows an N channel depletion-type insulated-gate field-effect transistor.

In the circuit shown, current flows from the negative side of the power supply through R_2 to the source of the transistor. It flows through the transistor to the drain and then through L_3 back to the positive side of the voltage source. The gate is connected to the negative automatic gain control voltage through L_2. The negative voltage applied to the gate will limit the width of the channel

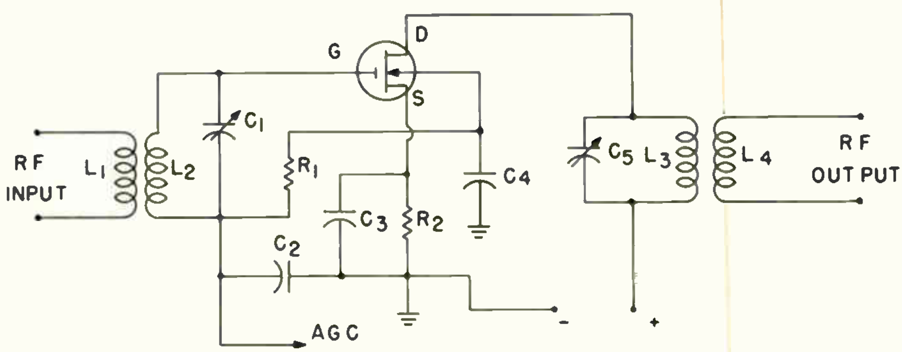


Fig. 17. An rf amplifier using an insulated-gate field-effect transistor.

and hence control the resistance of the channel.

In operation, the rf signal input is applied to L_1. This may be from another rf amplifier or directly from an antenna. The signal current flowing through L_1 induces a voltage in series with L_2. L_2 and C_1 form a series-resonant circuit. The resonant signal voltage is applied to the gate of the transistor and this voltage is applied in series with the negative agc voltage and hence varies the negative voltage on the gate at an rf rate. The varying signal voltage causes the resistance of the channel to vary and this causes the current flowing from the source through the transistor to the drain to vary. L_3 and C_5 form a parallel-resonant circuit. This high-impedance circuit develops a high signal voltage due to the varying current flowing through it. L_3 is inductively coupled to L_4 and the output from L_4 can be fed to another rf amplifier or to a mixer as it would be in the case of a superheterodyne receiver.

As in the case of the preceding circuit, a P-channel transistor could be used as well as an N-channel transistor. Also the enhancement type of insulated-gate transistors could be used. However, it is likely that most rf amplifiers using the insulated-gate field-effect transistors will be of the depletion type N channel transistors.

As mentioned previously, one of the disadvantages of the insulated-gate transistor is that they are easily damaged. Simply removing or inserting the transistor in the circuit when the voltages are applied could destroy the transistor due to high peak voltages built up in the gate circuit due to the very high resistance of the gate. Since the gate is actually insulated from the drain source by means of a layer of insulation the input resistance of the gate is extremely high. Pickup from a nearby power line can induce a high enough voltage in the gate to destroy the transistor if the gate circuit is open.

SELF TEST QUESTIONS

(k) What type of amplifiers are the rf amplifiers found in radio and television receivers?

(l) What type of resonant circuit is C_1 and L_2 in Fig. 10?

(m) What is the purpose of C_3 in the circuit shown in Fig. 10?

(n) What is permeability tuning?

(o) What is T_1 in the circuit shown in Fig. 13, and what purpose does it serve?

(p) Why are triode tubes often used as rf amplifiers in television receivers?

(q) Why does L_2, in Fig. 14, have fewer turns than L_1?

(r) What is the purpose of C_4 in the circuit shown in Fig. 14?

(s) In the circuit shown in Fig. 15 what are the combinations of L_1 and C_1 and L_2 and C_2 used for?

(t) Why must a negative voltage be applied to the gate of the junction field-effect transistor shown in Fig. 16?

(u) What is the primary advantage of the field-effect transistor in an rf amplifier circuit over the typical NPN or PNP transistor?

(v) Which of the two circuits, the one shown in Fig. 16 or the one shown in Fig. 17 has the least loading effect on the input circuit?

I-F Amplifiers

From the block diagram of the superheterodyne receiver which we have shown in Fig. 9, we see that the signal picked up by the antenna is amplified by an rf amplifier and then fed to a combination mixer and oscillator stage. In this stage the rf signal is mixed with a signal from a local oscillator. The local oscillator simply generates an rf signal which we use to mix with the incoming signal. The signal produced by the local oscillator is always a fixed frequency either above or below the frequency of the incoming signal. In most cases the oscillator is operated at a frequency above the frequency of the incoming signal.

Mixing the incoming rf signal with the locally generated signal produces two new signal frequencies in the mixer. It produces a signal equal to the sum of the frequency of the incoming signal plus the frequency of the oscillator and also a signal equal to the difference between the two frequencies. We use the difference signal as the intermediate frequency signal in a superheterodyne receiver. This signal is referred to as the i-f signal.

Since the frequency relationship between the incoming signal and the local oscillator is maintained constant, the i-f signal produced in the mixer will always have the same frequency. In other words, when the receiver is tuned to a signal of one frequency, for example a signal 1000 KHz, the oscillator might be operating at 1455 KHz. The difference between these two signal frequencies is 455 KHz and therefore this will be the i-f signal frequency. If we tune the receiver to a higher frequency, for example 1500 KHz, then the frequency of the local oscillator will also be increased so that when the rf and mixer stages are tuned to 1500 KHz, the local oscillator will be operating at 1955 KHz. Once again the difference frequency is 455 KHz. In other words, as long as we maintain the same difference between the incoming signal and the oscillator signal frequencies, the i-f signal frequency produced will be the same.

The advantage of a constant signal frequency for the intermediate frequency lies in the fact that we can design an amplifier with a higher gain when it is to be operated at a fixed frequency than we can if we have to be able to vary the frequency. Also, using a low frequency i-f gives us some advantages insofar as selectivity is concerned. We will go into this in detail later when you study superheterodyne receivers in more detail.

In this section of the lesson we are going to study some typical i-f amplifiers. Remember, even though we call the amplifiers intermediate frequency amplifiers, they are radio-frequency amplifiers because the signals they amplify are radio-frequency signals. However, they are different from the rf amplifiers you studied in the preceding section because they operate at a lower frequency, and also because once their frequency is set, it normally is not varied.

PENTODE I-F AMPLIFIERS

A typical pentode i-f amplifier is shown in Fig. 18. Notice that we have resonant circuits both in the input and the output of the stage, and that

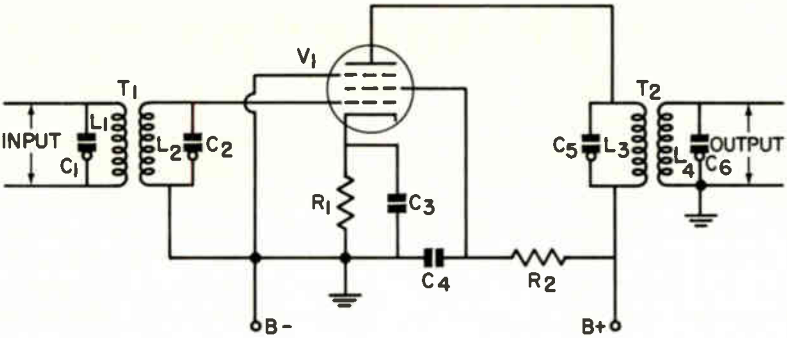


Fig. 18. A typical i-f amplifier using a pentode tube.

in many respects it resembles the pentode rf amplifier shown in Fig. 10.

In the circuit shown in Fig. 18, the input signal is obtained from the mixer output. In the output of the mixer we will have four signals. One signal will be equal to the frequency of the incoming signal, another signal will be equal to the frequency of the local oscillator. In addition, we have the two new signals produced by beating the incoming signal with the local oscillator signal. We have a signal equal to the sum of the two and a signal equal to the difference between the two. C_1 and L_1 form a parallel-resonant circuit which is resonant at a frequency equal to the difference between the local oscillator frequency and the incoming signal frequency. Thus, the parallel-resonant circuit acts as a high impedance at the difference signal frequency and a comparatively high signal voltage of this frequency is developed across it.

L_1 is inductively coupled to L_2. This causes a signal voltage to be induced in L_2. L_2 and C_2 form a series resonant circuit, and they also are tuned to the difference frequency, which from now on we will call the i-f frequency. In the series-resonant circuit we have a resonant voltage step-up, and this voltage is applied to the grid of V_1, and to the cathode through C_3. C_3, the cathode bypass is chosen so that it has a low reactance at the signal frequency.

The i-f signal applied between grid and cathode of V_1 causes the plate current flowing through the tube to vary and this varying plate current is fed to the parallel-resonant circuit consisting of L_3 and C_5. A high signal voltage is developed across the high impedance parallel-resonant circuit. L_3 is inductively coupled to L_4 and the combination of L_4 and C_6 form a series-resonant circuit which is also tuned to the i-f frequency. From the output, the voltage is fed to another i-f amplifier or to the second detector which separates the intelligence signal on the carrier from the carrier.

As in the rf amplifier the screen of the pentode tube is operated at signal ground potential by means of a bypass capacitor C_4 which has a low reactance at signal frequency. A positive dc voltage is applied to the screen of the tube through the resistor R_2. Normally the screen of the tube will be operated at a lower voltage than the plate and R_2 is chosen so that the voltage drop across it, when subtracted from the plate voltage, will provide the correct screen voltage for the tube.

In some circuits the suppressor

grid, the grid nearest the plate, is tied directly to the cathode of the tube. However, tieing the suppressor to ground will often introduce a small amount of degeneration in the stage and improve the stability of the stage. With modern pentode tubes there is such a high voltage gain in the stage, that even though the voltage fed from the plate of the tube through the tube, back to the grid of the tube is very small, it is possible that this voltage might be high enough to cause oscillation unless steps are taken to prevent it.

In some amplifiers you will find that the cathode of the tube is not bypassed. Leaving the cathode unbypassed again introduces some degeneration as it did in the case of audio amplifiers and this will tend to stabilize the stage and further prevent the stage from going into oscillation.

We mentioned earlier that the i-f amplifier is operated at a fixed frequency. Once the receiver is set up, the resonant frequency of the tuned circuits is not changed as you tune from station to station. When the receiver is set up, the various resonant circuits are all adjusted to resonate at the i-f frequency. This is usually accomplished by feeding a signal from an instrument called a signal generator into the i-f amplifier. The signal generator is set to the intermediate frequency and then the various resonant circuits are adjusted for maximum output from the i-f amplifier. The coupling between the coils is adjusted by the manufacturer so that when the various circuits are tuned for maximum output the required bandpass will be obtained. It is possible to do this in radio receivers because the bandwidth required in radio reception is comparatively narrow. In television receivers, however, the i-f amplifier presents special problems and alignment methods other than peaking the transformers or coils at one frequency must be used.

In the early days of radio, triode i-f amplifiers were used, but since the development of the pentode tube, the triode tube has not been used for this purpose. You can get a higher gain with a pentode i-f amplifier than with a triode, and in addition the pentode does not require neutralization.

TRANSISTOR AMPLIFIERS

A typical transistor i-f amplifier is shown in Fig. 19. Notice that this circuit is practically identical to the

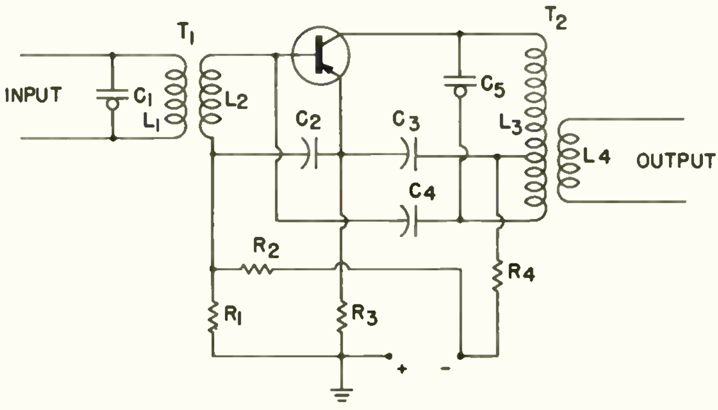


Fig. 19. A typical transistor i-f amplifier using a PNP transistor.

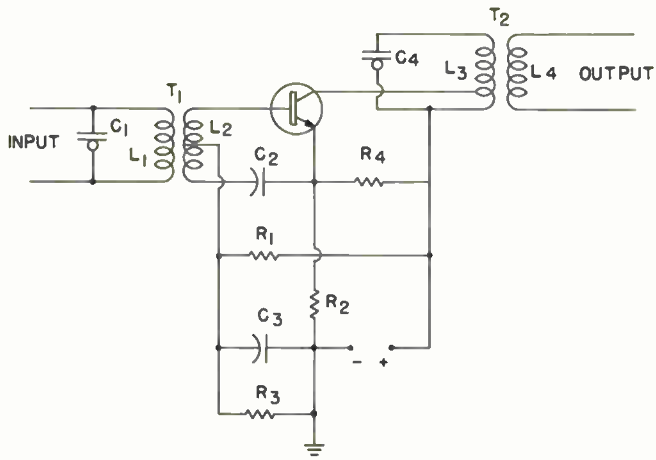


Fig. 20. Transistor i-f amplifier using NPN transistor.

circuit of the rf amplifier shown in Fig. 14.

In Fig. 19, C_1 and L_1 form a parallel-resonant circuit and would be in the collector circuit of the mixer stage. L_1 is inductively coupled to L_2. L_2 has fewer turns than L_1 so that we have a step-down transformer to match the low input resistance of the transistor circuit to the high resistance of the parallel circuit. The signal is applied to the base of the transistor and to the emitter through C_2. The i-f signal current from the collector is fed to the parallel resonant circuit which consists of L_3 and C_5. L_4, which has fewer turns than L_3, is inductively coupled to L_3 so that a signal current will be induced in L_4. The output from L_4 would probably be connected to another transistor i-f amplifier or to a detector stage. As in the rf stage shown earlier, C_4 provides neutralization and feeds a signal back into the base which is 180° out-of-phase with the signal fed from the collector to the base through the transistor.

Forward bias for the transistor is provided by the voltage divider network consisting of R_1 and R_2. As in the previous circuits, R_3 is put in the circuit for bias stabilization, to prevent thermal runaway of the transistor.

Another transistor i-f amplifier is shown in Fig. 20. This amplifier uses an NPN transistor whereas the one shown in Fig. 19 uses the PNP transistor. In addition to the different transistor types, the method of obtaining neutralization is somewhat different. Notice the resistor in the emitter circuit, R_2. This resistor is not bypassed and therefore a signal voltage will be developed across this resistor. This signal voltage is fed through C_2 into the lower end of the coil L_2. The center tap of L_2 is at signal ground potential, because C_3 has a low reactance at the signal frequency. The lower end of L_2 is inductively coupled to the upper end so that a voltage is induced in the upper end of the coil 180° out-of-phase with the signal fed to the lower end through C_2. This voltage is fed to the base and will neutralize any signal voltage fed from the collector back into the base through the transistor itself.

Notice that the collector is connected to a tap on the coil L_3. The output resistance of the transistor is comparatively low and by feeding it to a tap on the coil in this manner we prevent loading of the parallel resonant circuit made up of L_3 and C_4. This prevents loading of the resonant circuit which in turn would cause a reduction in the selectivity in the circuit.

Field-effect transistors can also be used as i-f amplifiers, but since they are considerably more expensive than the PNP and NPN types they have not been widely used in this application.

VIDEO I-F AMPLIFIERS

Television receivers are superheterodyne receivers just like radio receivers. However, the rf, mixer and oscillator stages operate at higher frequencies than regular receivers designed for reception on the standard broadcast band.

The i-f amplifier used to amplify the picture signals is called a video i-f amplifier. A video i-f amplifier differs from a sound i-f amplifier inasmuch as it operates at a much higher frequency and must have a much wider bandwidth. This is due to the fact that a wide range of signals is required in order to reproduce a TV picture. Some large areas may be reproduced by comparatively low frequency signals, but the fine detail in a picture is reproduced by comparatively high-frequency signals. The video i-f amplifier must be able to pass the video i-f carrier and the sidebands which contain the video signal. Also, in color TV receivers they must pass an additional signal called the color subcarrier. This signal carries the color information in the picture. In addition, modern superheterodyne TV receivers must be capable of passing the sound i-f signal through at least part of the video i-f amplifier. The sound i-f signal differs from the picture i-f signal by 4.5 megahertz. Therefore the video i-f amplifier must be capable of at least passing with some amplification the sound i-f signal which will be 4.5 MHz lower in frequency than the video i-f carrier frequency.

You'll study video i-f amplifiers later, but here we want you to get some general idea of what the circuit looks like and the problems involved.

A typical video i-f amplifier using a pentode tube is shown in Fig. 21.

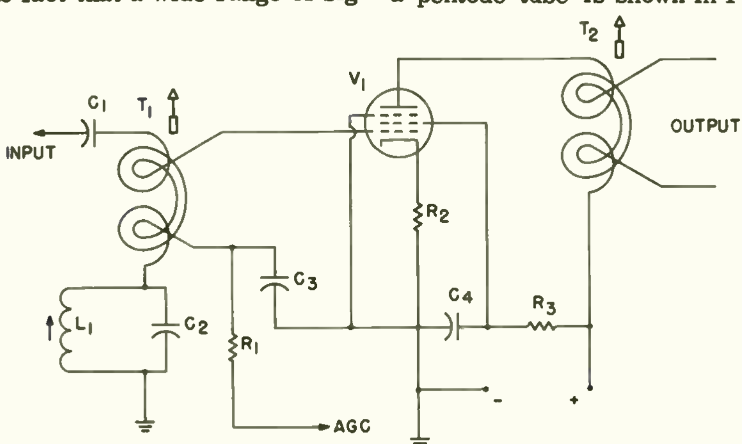


Fig. 21. A pentode video i-f amplifier.

This is typical of the circuits you will find in the first video i-f amplifier of a color TV receiver or of a black and white TV receiver. The input signal is fed to the i-f amplifier from the mixer. It is fed through the capacitor C_1 to the primary of T_1. T_1 is the input i-f transformer and the coils are bifilar wound. This means that the coupling between the coils is extremely tight. The arrow above the two coils indicates that there is an iron slug inside the coils which can be adjusted to tune the two coils to a particular frequency.

An example of how a bifilar wound transformer is constructed was shown in Fig. 8. Notice that the two windings are interlaced to provide very tight coupling between the primary and secondary windings of the transformer.

The combination of L_1 and C_2 form a parallel-resonant circuit. This circuit is resonant at a frequency of 47.25 MHz. This parallel-resonant circuit is referred to as an adjacent channel sound trap. When your TV receiver is tuned to channel 4, for example, if you happen to be near enough to pick up a signal from another station operating on channel 3, the sound signal from the channel 3 station might get through the rf and mixer stages and could cause some interference in the picture. By inserting the parallel-resonant circuit in series with the primary of T_1, if there is any signal current from the channel 3 sound flowing in the circuit, almost all will be dropped across the high impedance of this parallel-resonant circuit. Very little will be fed from the primary to the secondary of T_1.

The grid of the tube is fed to the automatic gain control circuit through R_1. The automatic gain control will provide a negative voltage for the grid of the tube. This voltage will vary in amplitude depending upon the strength of the signal received. If the signal is very strong, a rather high negative voltage will be fed to the grid of V_1. This will reduce the gain of the stage and prevent overloading in following stages. On the other hand, if the signal is weak, the negative voltage fed to the grid of V_1 will be quite low so that the tube will operate at its maximum gain.

T_2 is another bifilar wound i-f transformer. Again, the primary and secondary coils are very tightly coupled.

You might notice that there is no capacitor in the plate circuit of V_1. In spite of this, we have a parallel-resonant circuit in the plate circuit of V_1. You might wonder how this is so, but remember that there is a certain capacity in the tube itself. There is capacity between the plate and cathode and this is in effect electrically across the primary of T_2. This capacity along with the wiring capacity in the circuit is all the capacity we need at the frequencies involved to form a parallel-resonant circuit along with the primary winding of T_2. Video i-f amplifiers usually operate in the 40 MHz region, and at these high frequencies only a small amount of capacity is needed in a resonant circuit.

A typical transistor video i-f amplifier is shown in Fig. 22. This i-f amplifier is the second video i-f amplifier taken from a portable TV receiver. The input signal is received from the first video i-f amplifier and coupled from L_1 to L_2. L_2 has fewer turns than L_1 in order to match the transistor input circuit. C_1 is placed in the circuit to prevent L_2 from shorting out the forward bias across the emitter-base junction of the transistor. The for-

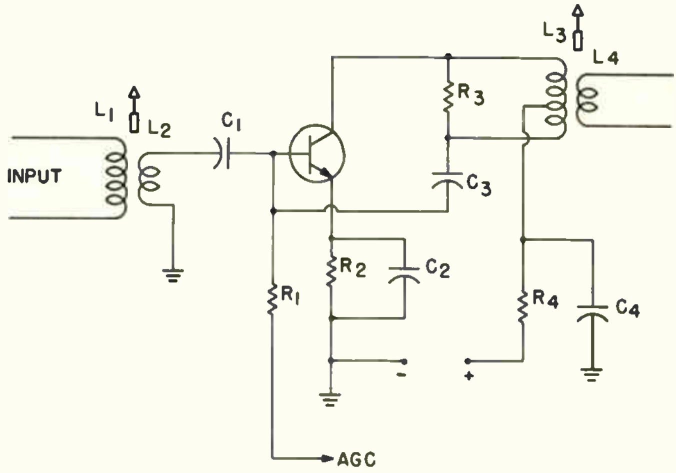


Fig. 22. A transistor video i-f amplifier.

ward bias is applied to the base through R_1. Again, this bias is obtained from the agc system, and is varied to vary the gain of the transistor depending upon the strength of the signal being received.

The coil L_3, which is in the collector circuit of the transistor, along with the transistor capacity and the wiring capacity, form a parallel-resonant circuit. The arrow of the coil indicates that the circuit can be tuned by means of a slug which moves in and out of L_3. Notice that the collector voltage is fed to a center tap on the coil. The center tap is held essentially at ground potential by the capacitor C_4. A voltage is induced in the lower half of L_3 which is 180° out-of-phase with the voltage in the upper half. The voltage from the lower half is fed through C_3 back into the base circuit in order to provide neutralization and to cancel out any signal fed through the transistor from the collector to the base.

The resistor R_3, which is connected across L_3, is a loading resistor. The purpose of this resistor is to load the resonant circuit in order to help get the broad frequency response required of a video i-f amplifier.

L_4 is inductively coupled to L_3 and feeds the next video i-f amplifier.

In aligning a video i-f amplifier such as shown in Fig. 22, L_1 is not tuned to the same frequency as L_3. By tuning the resonant circuits to different frequencies we broaden the response of the amplifier. For example, the resonant circuit in which L_1 is located may be tuned as shown by curve A in Fig. 23. At the same time L_3 and its circuit might be tuned as shown in curve B. The overall response produced by the two amplifiers may then look like curve

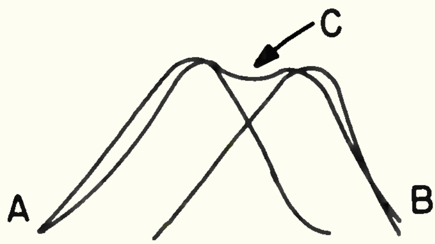


Fig. 23 Response of curves A and B together may give overall response like C.

C which has a wider bandwidth than either curve A or curve B. You will study video i-f amplifiers in detail later and learn more about how the various resonant circuits are adjusted for resonance.

SELF TEST QUESTIONS

(w) What determines the i-f frequency in a receiver?

(x) Why can we usually operate an i-f amplifier at higher gain than an rf amplifier?

(y) Why is the cathode resistor in some pentode i-f amplifiers left unbypassed?

(z) In the circuit shown in Fig. 20, why is the collector connected to a tap on L_3?

(aa) Where is the neutralization for the transistor in the circuit shown in Fig. 20 obtained?

(ab) What is the chief difference between the video i-f amplifier in a television receiver and the sound i-f amplifier in a radio receiver?

(ac) Why is there no capacitor across the primary winding of T_2 in the circuit shown in Fig. 21?

(ad) What purpose does R_3 in Fig. 22 serve?

(ae) What purpose does C_3 in Fig. 22 serve?

Radio-Frequency Power Amplifiers

So far the rf amplifiers we have been discussing are the types found in radio and television receiving equipment. They are voltage amplifiers. They are used to build the weak signal voltage to a reasonably high value before the signal is fed to a detector to extract the intelligence that has been used to modulate the carrier. In radio and television transmitters and in other industrial applications rf power amplifiers are used. RF power amplifiers differ from voltage amplifiers in a number of respects. Not only are the tubes generally much larger than those used in voltage amplifiers, but also the circuit configurations differ in several respects from those used in voltage amplifiers.

A TRIODE POWER AMPLIFIER

Although triode tubes are not very frequently used as rf voltage amplifiers except in the vhf region, they are quite widely used as power amplifiers. Many high-power transmitters today use triode rf power amplifiers.

Schematic diagrams of two triode rf power amplifiers are shown in Fig. 24. Since there are several differences between these two circuits we'll look at the circuits one at a time.

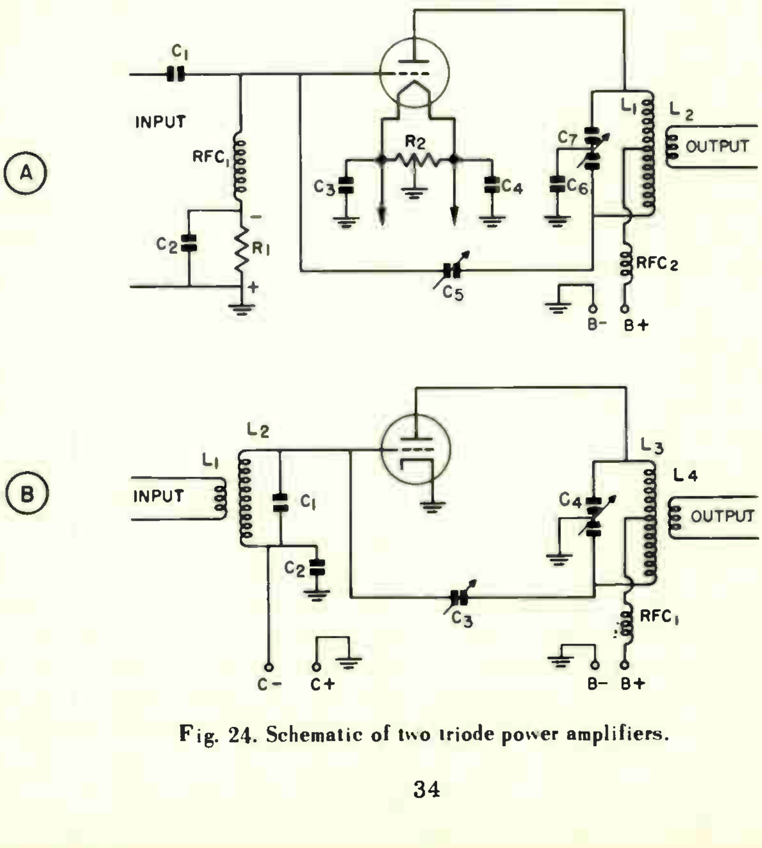


Fig. 24. Schematic of two triode power amplifiers.

In the circuit shown in Fig. 24A we have shown a filament-type tube. This type of tube is frequently found in transmitting applications because a solid tungsten filament or a thoriated tungsten filament will stand up better than an oxide-coated cathode at the high voltages usually used in transmitting tubes. In the circuit shown in Fig. 24B we have shown a cathode-type triode tube so that you will see the difference between the two types.

Examining the grid circuit of the diagram shown in Fig. 24A, you'll see that connected between the grid of the tube and ground we have a coil called a radio-frequency choke. This is abbreviated RFC on the diagram. Also connected between this choke and ground we have a resistor R_1. Power amplifiers are normally operated as class C amplifiers. This means that the input signal drives the grid of the tube positive so that electrons flow from the filament of the tube to the grid. These electrons will strike the grid and then flow from the grid through the rf choke to the parallel combination of R_1 and C_2. Some of the electrons will flow through R_1 to ground, but most of them will charge C_2. During the portion of the cycle when the grid is not drawing current, C_2 will discharge through R_1. The time constant of R_1 and C_2 is selected so that C_2 does not appreciably discharge between cycles of grid current. R_1 is selected so that the average grid current provides the bias required for the tube. Usually for class C operation, this is somewhere between two and four times cut-off bias voltage.

Notice the bypass capacitors and the tapped resistor in the filament circuit. Each side of the filament is bypassed to ground, and a tapped resistor is connected across the filament, and the center tap of this resistor is grounded. Sometimes instead of using a tapped resistor of this type, the filament winding of the transformer used to supply the voltage to heat the tube is center-tapped and the center tap is grounded.

The filament of a transmitting tube is rather heavy, and as a result can be operated from ac. The temperature of the filament changes so slowly that there is no heating and cooling of the filament as the ac voltage goes through its cycle.

With the high bias on the tube, plate current will flow through the tube in a series of pulses. There will be a pulse of plate current each time the input signal drives the grid sufficiently in a positive direction to overcome the bias. Because the grid is driven positive, the peak current reached during each pulse will be high, but because the operating bias on the tube will be substantially beyond cutoff bias, plate current will flow for less than one half of each cycle.

In the plate circuit of this stage, we have a parallel resonant circuit. Notice that the coil L_1 is tapped, and the center tap is connected through another radio-frequency choke (usually called an rf choke) to B+. The B+ voltage applied to the plate of the tube is thus applied through the choke and half of L_1. Capacitor C_7 is a variable capacitor called a split stator capacitor. You will remember that a variable capacitor has two separate plates, one set that is stationary, called the stator, and the other set that rotates, called the rotor. In a split-stator type capacitor there are two sets of stator plates, which are insulated from each other. The rotor consists of two separate sets of plates, one for each set of stator plates, but the rotor plates are electrically connected together. The rotor is operated at signal ground potential by grounding it through C_6 and the stators are connected to the ends of L_1.

The parallel resonant circuit in the plate circuit of the tube is more or less shock-excited by the pulses of plate current received from the tube. The pulse of the plate current sets up circulating currents in the tank circuit. These pulses from the tube are more or less smoothed out by the tank circuit so that the current circulating back and forth in the resonant circuit is a sine wave.

Since the tube used in this circuit is a triode tube, there will be energy fed from the plate of the tube back to the grid circuit. Because a power amplifier develops considerable power in the plate circuit, there will be enough energy fed back into the grid circuit to result in oscillation. This oscillation is overcome by feeding a signal from the tank circuit through C_5 back into the grid circuit of the tube. Notice that the plate of the tube connects to one end of the tank circuit and C_5 connects to the other end. The voltage at the two ends of the tank circuit will be of opposite polarity. Therefore, the signal fed through C_5 back into the grid circuit is of opposite polarity to the signal fed from the plate to the grid circuit through the tube capacity. C_5 is usually an adjustable capacitor that can be adjusted to feed exactly the same amount of signal into the grid circuit as is fed through the tube capacity. C_5 is called a neutralizing capacitor. It is used to feed back energy into the grid circuit to neutralize or cancel the energy fed back into the grid circuit through the tube.

The output signal is taken from the tank circuit by inductively coupling another coil, marked L_2 on the diagram, to L_1.

When a tube is used in a circuit of this type where its bias is developed by the grid current flowing through the grid resistor, we say the tube is self-biased. This type of bias is entirely satisfactory when the stage is operating properly and when the normal signal drive is applied to the input. However, if a defect develops in a preceding stage so that no signal is applied to the input, the grid of the tube will not be driven positive and there will be no current flow. Capacitor C_2 will discharge through R_1 so that the bias applied to the tube will disappear. Once the bias disappears there will be a very high current flow from the filament to the plate of the tube and unless there is some safety measure incorporated in the circuit so that the circuit will be opened when the current goes beyond a certain level, the current will rise to such a high value that the tube will be destroyed.

In the circuit shown in Fig. 24B we have tuned circuits in the input and output circuits. Coil L_1 is inductively coupled to L_2 and its energy fed into L_2. L_2 and C_1 form a series resonant circuit. Here, instead of using self-bias, a fixed bias voltage is applied to the grid circuit of the tube. Notice that the negative terminal of the C battery or C bias supply connects to the coil, and through the coil to the grid. The positive terminal of the C supply is grounded. Sometimes you'll find the bias connections on the diagram labeled C- and C+ as in Fig. 24B and on other occasions you'll find them labeled bias.

The resonant circuit in the plate circuit of this stage is the same as the circuit shown in Fig. 24A. Again, the stage is neutralized by feeding energy from the tank circuit through a capacitor (C_3) back into the grid circuit of the tube.

In some triode power amplifiers you'll find a combination of fixed bias and self bias. In this type of circuit the two types of bias shown in Fig. 24 are incorporated into a single circuit. This is done by selecting a value of resistor in the grid circuit somewhat smaller than what is needed to develop the full bias required by the tube. The remainder of the bias is supplied by a separate bias power supply. For example, if a tube requires a bias of 100 volts for class C operation, a resistor might be put in the grid circuit that would develop 50 volts bias and then a fixed power supply used to provide the other 50 volts. The advantage of this arrangement is that the bias is somewhat self-regulating. In other words if the grid current increases, the bias will increase and tend to prevent over-driving the tube. At the same time, the fixed bias would also protect the tube if the preceding stage develops a defect so that the signal drive is not supplied to the grid of the tube.

TETRODE POWER AMPLIFIERS

Tetrode power amplifiers have two big advantages over the rf triode power amplifiers. Perhaps the greater advantage lies in the fact that a tetrode has a very high power sensitivity. This means that you need only a small input signal to drive the tube hard enough to produce a rather large signal in the output. Thus the power gain in a tetrode rf power amplifier is much greater than in a triode power amplifier, because a substantial amount of driving power is usually required to drive a triode.

The second big advantage of a tetrode power amplifier is the fact that the screen grid effectively shields the grid of the tube from the plate. With careful circuit design it is usually possible to design a circuit that does not require neutralization. This is a big advantage if a power amplifier is to be used over a wide frequency range. In a triode power amplifier, if you have occasion to change the frequency at which the amplifier is operating, you often have to readjust the setting of the neutralizing capacitor. However if a tetrode stage can be operated without neutralization, this problem is not encountered. Of course, in a radio broadcast transmitter that is designed to operate on one specific frequency, neutralizing the triode doesn't present any great problem, and once it is neutralized you do not have to re-neutralize it unless you change the tube. However, in many communications applications, it is necessary to have transmitters that can be operated on a number of different frequencies. In such a case, a tetrode tube that does not need neutralization is quite advantageous.

The tetrode tubes used in modern power amplifiers are beam power tubes. You will remember that a beam power tube is a tube made so that the cathode emits electrons in two streams or beams from opposite sides of the cathode. The electrons are further focused into a beam by means of beam-forming plates which are connected inside the tube to the cathode of the tube. Although you might actually consider the beam-forming plates as separate elements, the general practice is to ignore them when counting the tube elements so the tube is considered to be a tetrode or four-element tube.

A schematic diagram of a tetrode rf power amplifier is shown in Fig. 25. Notice that we have a resonant circuit in the input. Again, this tube circuit uses self bias; the bias needed to operate the tube in the Class C amplifier is developed

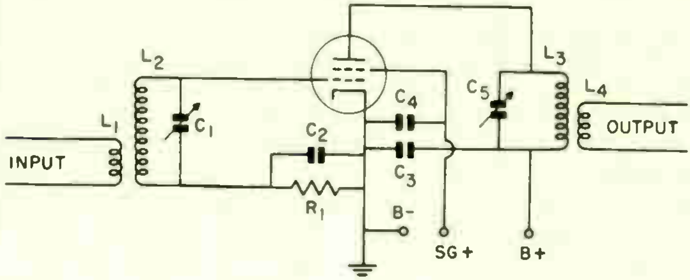


Fig. 25. A tetrode rf power amplifier.

across the parallel combination of R_1 and C_2 by the grid current. Again, grid current flows because on the positive half cycle the signal drives the grid positive and electrons are attracted to the grid.

A parallel resonant circuit is used in the plate circuit, and the output is taken by inductively coupling L_4 to L_3.

The screen grid of the tube is operated with a positive voltage applied to it. Sometimes this voltage is obtained from a separate power supply or from a tap on the main power supply. Sometimes a dropping resistor is connected between the plate supply and the screen grid of the tube in order to drop the screen voltage to a value lower than the plate voltage.

A push-pull tetrode power amplifier is shown in Fig. 26. Here notice that the circuit configuration is quite similar to the push-pull circuits used in audio work with the exception of the resonant circuits used in the grid and plate circuits of the tube. Notice that split stator tuning capacitors have been used in both the input and the output circuits. Again, because the feedback from the plate of the tube to the grid of the tube is low due to the shielding effect of the screen, neutralization is not required. However, sometimes you will find tetrode power amplifiers using neutralization.

Pentode rf power amplifiers are sometimes used, but they are not nearly as common as tetrode power amplifiers. The circuits used for

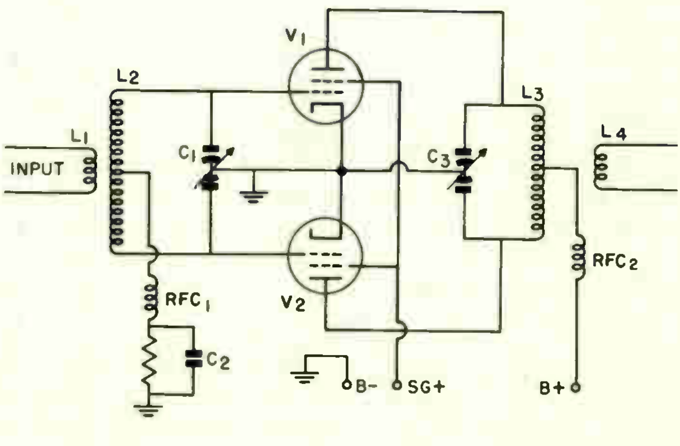


Fig. 26. A push-pull tetrode rf power amplifier.

pentode tubes are similar to those used for tetrodes. The suppressor grid of the pentode tube is usually connected to B-.

TRANSISTOR POWER AMPLIFIERS

A few years ago there were relatively few transistor rf power amplifiers because the rf power that could be generated by transistors was quite limited. However, a great deal of progress has been made in transistor design and manufacture and now there are many transistor rf power amplifiers in use. Of course, there are no transistors available that can develop the very high rf powers that vacuum tubes can develop, but low and medium power rf transistors are available.

Transistor power amplifiers may be used in either the common-emitter or common-base circuits. The common-emitter circuit is more stable, but at very high frequencies the emitter lead inductance may restrict the power capability of the transistor. In this case the common-base circuit may provide a higher power gain, but it will be more unstable than the common-emitter circuit.

Transistor rf power amplifiers may be operated as Class A, Class B or Class C amplifiers. Class A amplifiers provide extremely good linearity but the efficiency is low. This type of power amplifier is used only when the power requirements are quite low. A Class B power amplifier will provide a higher power gain and better efficiency. Of course, a Class C power amplifier will provide the best efficiency, but the harmonic output from a Class C power amplifier will be quite high. The tank circuits used with Class C power amplifiers are designed to offer a high impedance to the harmonics and a low impedance to the fundamental frequency.

Biasing Methods.

For Class A bias on a transistor power amplifier, we must have a forward bias across the emitter-base junction as in the case of transistor voltage amplifiers. For Class B bias, the bias across the emitter-base junction of the transistor is zero. For Class C bias, we must have a reverse bias across the emitter-base junction so that current will flow through the transistor only on the peak of the rf cycle. That overcomes the reverse bias and drives the emitter-base junction into the conduction region. As in the case of vacuum-tube Class C amplifiers, current will flow through the transistor for less than half a cycle.

Two typical circuits for developing Class C bias for a transistor power amplifier are shown in Fig. 27. In the circuit shown in Fig. 27A, the incoming rf signal drives the

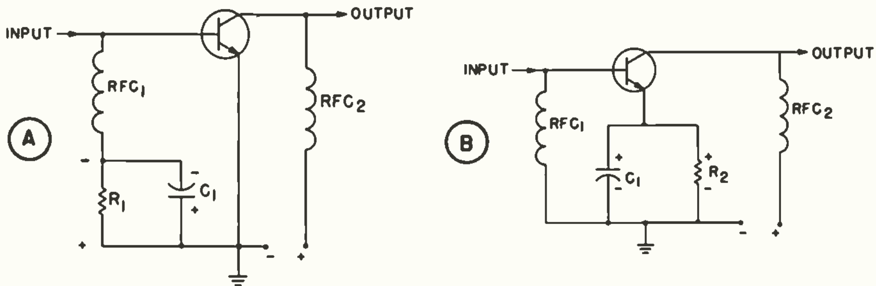


Fig. 27. Two methods of producing bias for Class C power amplifiers.

base of the transistor positive and electrons leave the base and flow through the rf choke and charge the capacitor C_1 with the polarity shown. Some of the electrons flow through R_1 to ground developing a voltage drop across this resistor having the polarity shown.

During the portion of the rf cycle when the transistor is not conducting, C_1 discharges through R_1 maintaining the reverse bias across the emitter-base junction essentially constant.

Another method of obtaining Class C bias is shown in Fig. 27B. In this circuit, when the input signal drives the base positive, current flows through the resistor R_2 to the emitter, across the emitter-base junction of the transistor, across the base and the base-collector junction and then through the rf choke RFC_2 back to B+. The electrons flowing through R_2 will charge capacitor C_1 with the polarity shown. During the portion of the cycle when the base is not driven positive by the rf signal, C_1 will discharge through R_2 maintaining the emitter positive with respect to ground. Thus the emitter will be positive with respect to the base and will have a reverse bias across the emitter-base junction so there will be no current flow through the transistor.

Typical Circuits.

A typical Class C transistor rf power amplifier is shown in Fig. 28. In this circuit the biasing method shown in Fig. 27B is used to provide a reverse bias across the emitter-base junction.

The rf signal is fed into the input and on the positive half of the cycle the rf signal overcomes the reverse bias across the emitter-base junction and electrons flow through the transistor. The choke in the collector circuit completes the dc current path through the transistor. The rf signal is fed to the output network consisting of C_2, C_3, L_1 and L_2. C_2 is adjusted for resonance and C_3 is adjusted to obtain the desired loading.

When more power is required from an rf power amplifier than can be obtained from a single transistor, two or more transistors can be used either in push-pull or in parallel. In push-pull operation transformers must be used for proper input signal phase. This works out quite satisfactorily at lower i-f frequencies, but at very high frequencies it is difficult to build transformers which provide the required impedance

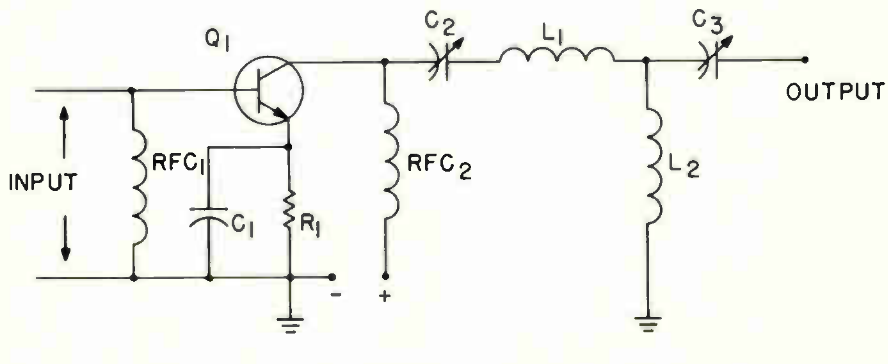


Fig. 28. A Class C transistor power amplifier.

transfer. Therefore parallel operation of transistors is generally preferred at vhf over push-pull operation.

In Fig. 29 we have shown a schematic diagram of two transistors operated in parallel. If additional power is required, a third parallel transistor in an essentially identical circuit could be added.

Previously we mentioned the effect of emitter-lead inductance at vhf. The effect of this inductance is tuned out by the capacitors C_3 and C_4.

In a circuit of this type, it is desirable to have transistors with matched characteristics in order to insure that each transistor will pick up half of the load.

To check the operation of the transistors to be sure that each is handling half the total current, we measure the voltage across R_1 and the voltage across R_2. The voltages across these two resistors should be equal; this would indicate each transistor is picking up half of the load.

SELF TEST QUESTIONS

(af) What is the purpose of C_5 in Fig. 24A?

(ag) Why is it possible to operate the filament of the triode tube shown in Fig. 24A from ac power?

(ah) How are the pulses of current that flow through a Class C amplifier stage smoothed into a sine wave?

(ai) What are the two big advantages of the tetrode power amplifier over the triode power amplifier?

(aj) What type of bias is placed across the emitter-base junc-

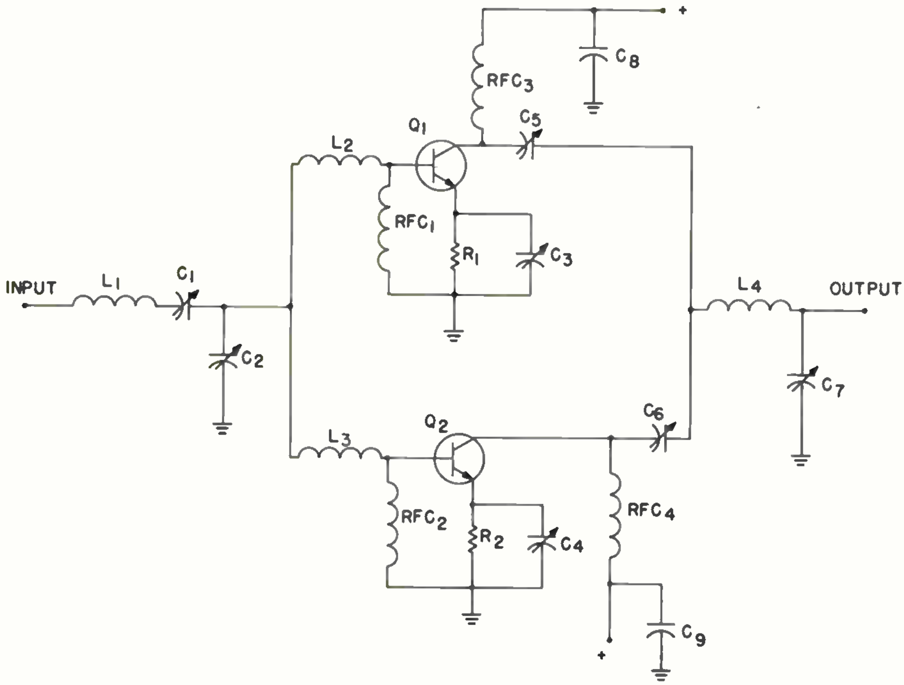


Fig. 29. A high-frequency power amplifier using two transistors in parallel.

tion of a transistor operated as a Class C power amplifier?

(ak) Across what parts is Class C bias developed in the circuit shown in Fig. 27A?

(al) What purpose do the capacitors C_3 and C_4 in the circuit shown in Fig. 29 serve?

ANSWERS TO SELF TEST QUESTIONS

(a) The basic difference is in the load found in the rf amplifier. It is usually some type of resonant circuit.

(b) The basic difference is how the voltage is applied to the coil and capacitor in the resonant circuit. If the voltage is applied to the coil and capacitor in series, it is a series-resonant circuit, but if it is applied to the coil and capacitor in parallel, it is a parallel-resonant circuit.

(c) Sideband frequencies are frequencies above and below the carrier frequency produced when we add the intelligence to be transmitted to the carrier frequency.

(d) The bandwidth of a resonant circuit is the band of frequencies that will be passed or amplified by the circuit with a gain equal to at least 70.7% of the gain obtained at the resonant frequency of the circuit.

(e) On the L-C ratio of the resonant circuit and on the Q of the resonant circuit.

(f) By loading the circuit with a suitable resistance.

(g) Critical coupling is the point at which maximum signal transfer occurs from one coil to another. This will occur when all the flux lines produced by one coil cut the turns of the other coil.

(h) When we speak of over-coupling we mean that the coils are coupled together beyond the critical coupling point. It results in some drop in output and produces a so-called double-hump response curve.

(i) The moisture and dust particles will introduce resistance into the circuit and lower the Q of the circuit. This will tend to reduce the output from the resonant circuit and also reduce the selectivity.

(j) At the high frequencies used in the rf section of a television receiver, moving a wire may change the inductance or capacitance in a resonant circuit sufficiently to upset the performance of that circuit. This is particularly true of the coils and leads used in conjunction with the high vhf channels in a TV tuner. It is even more true of the parts and leads in a uhf tuner.

(k) The rf amplifiers in radio and television receivers are Class A voltage amplifiers.

(l) C_1 and L_2 in Fig. 10 form a series-resonant circuit.

(m) C_3 is a screen bypass capacitor. It is chosen so that it has a low reactance at the signal frequencies amplified by the stage. Therefore as far as the screen is concerned, it is essentially connected to ground at signal frequencies. This permits the screen to act as an effective shield between the plate and grid of the tube, to prevent feedback from the plate to grid that could cause oscillation.

(n) Permeability tuning is a tuning system where the inductance of the coils is varied by moving a powdered iron slug in and out of the coil. The resonant frequency of the circuit is varied by varying the inductance of the coil rather than by changing the capacity of the tuning capacitor.

(o) T_1 is a balun. It is used to match the balanced transmission line, which has an impedance of 300 ohms, to the unbalanced input of the rf stage, which has an input impedance of 75 ohms. Thus the transformer matches a balanced circuit to an unbalanced circuit, at the same time it matches a 300-ohm circuit to a 75-ohm circuit.

(p) Triode tubes generate a lower internal noise than pentode tubes. Low noise level is extremely important in the rf stage of TV receivers, particularly in the reception of the high-band vhf TV channels.

(q) The combination of L_1 and L_2 form a transformer because L_1 is inductively coupled to L_2. L_1 along with C_1 form a series-resonant circuit. A step-down transformer is used to feed the signal to the input of the transistor because the transistor has a comparatively low input resistance. When using a step-down transformer (having fewer turns on L_2 than on L_1), loading of the resonant circuit can be kept to a minimum. Excessive loading of the resonant circuit would reduce the selectivity of the resonant circuit.

(r) C_4 is used to feed a small amount of signal from the output back into the base of the transistor. The energy fed through C_4 is 180° out-of-phase with the signal fed from the collector back to the base through the transistor itself. The signal fed through C_4 cancels the signal fed internally through the transistor. C_4 is a neutralizing capacitor and prevents the stage from going into oscillation.

(s) L_1 and C_1 form a parallel-resonant circuit. They prevent undesired signals in the sound i-f frequency region from getting to the rf amplifier. C_2 and L_2 form a series-resonant circuit. Since they have a low impedance they bypass the signals around 45 MHz. 45.75 MHz is the picture i-f signal frequency. The trap prevents interfering signals from stations operating at or near this frequency from reaching the rf amplifier.

(t) A negative voltage is applied to the gate to prevent electrons from flowing to the gate. This would cause current to flow in the gate circuit and lower the input resistance of the transistor.

(u) The field-effect transistor has a very high input resistance. Therefore the transistor has very little or no loading effect on the resonant circuit which permits a high degree of selectivity. In addition, very high gains are possible with the field-effect transistor.

(v) The circuit shown in Fig. 17. This circuit makes use of an insulated-gate field-effect transistor. There will be no current flow in the circuit. In the circuit shown in Fig. 16 there will be some small reverse current which has some slight effect on the input resistance of the transistor.

(w) The i-f frequency is determined by the difference between the frequency of the incoming signal and the frequency of the local oscillator.

(x) I-F amplifiers operate at a fixed frequency and therefore the circuit is usually somewhat simpler and can be better shielded than an rf amplifier which must be adjustable over a frequency range.

(y) To introduce a small amount of degeneration which may prevent the stage from going into oscillation.

(z) The collector is connected to a tap to prevent the comparatively low output resistance of the transistor from loading the parallel-resonant circuit made up of L_3 and C_4.

(aa) Neutralization is obtained by taking a signal from the emitter and feeding it through C_2 into L_2. L_2 inverts the signal so that the signal fed to the base will be out-of-phase with the signal fed back to the base through the collector-to-base capacity. The emitter is left unbypassed so that a signal voltage for neutralization will be developed across R_2.

(ab) The video i-f amplifier in a television receiver will have a much wider bandwidth than the sound i-f amplifier in a radio receiver. In order to get this wider bandwidth, the video i-f amplifier is operated at a much higher frequency.

(ac) A capacitor is not necessary at the high frequencies used in video i-f amplifiers. There is already enough capacity in the tube and in the circuit wiring to provide the capacity necessary to form a parallel-resonant circuit along with the primary winding of T_2.

(ad) R_3 is used to load the parallel-resonant circuit consisting of L_3 plus the transistor and distributed capacity in the circuit. The resonant circuit is loaded giving wider bandwidth.

(ae) C_3 is the neutralizing capacitor. A signal voltage is fed through it to neutralize the voltage fed through the collector–to–base capacity of the transistor.

(af) C_5 is a neutralizing capacitor. It is used to feed energy from the plate circuit back into the grid circuit to cancel out the energy fed from the plate circuit to the grid circuit through the interelectrode capacity of the tube.

(ag) The filament of the transmitting tube is quite heavy and does very little heating or cooling as the ac current goes through its cycle. Operating the filament on ac does not introduce hum into the circuit.

(ah) The tank circuit consisting of the coil and the capacitor in the plate circuit of the rf power amplifier smooth the current pulses from the tube into a pure sine wave.

(ai) The tetrode tube provides a higher power gain, and a properly designed tetrode amplifier usually does not require neutralization.

(aj) The emitter-base junction of a transistor Class C power amplifier is reverse biased.

(ak) R_1 and C_1.

(al) They are used to tune out the emitter lead inductance.

Lesson Questions

Be sure to number your Answer Sheet B203.

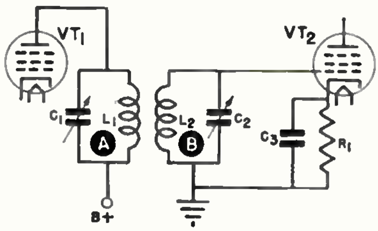
Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible, so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What type of load will be found in the plate circuit of most rf amplifiers?

2. If a 1000-KHz signal is modulated by a 3000-cycle audio signal, what will the frequencies of the sidebands be?

3. In the circuit shown to the right what type of resonant circuit is the circuit marked A; what type is the resonant circuit marked B?



4. Which type of rf amplifiers are found in radio and TV receivers, voltage amplifiers or power amplifiers?

5. Why are triode rf amplifiers frequently used in television receivers?

6. What is the purpose of C_4 in the circuit shown in Fig. 14?

7. What purpose do L_1 and C_1 in the circuit shown in Fig. 15 serve?

8. What characteristic of the junction-type field-effect transistor makes it ideally suited for use as an rf amplifier?

9. Why is the suppressor grid of a pentode tube, used as an i-f amplifier, often connected to ground rather than the cathode of the tube?

10. How is bias developed for the Class C transistor power amplifier shown in Fig. 28?

FEAR LEADS TO FAILURE

No matter how hard a person may work for success, there is nothing which can help him if he is always doubting his own ability-if he is always thinking about failure.

To be ambitious for wealth yet always expecting to be poor is like trying to get past a vicious dog when afraid of the dog and uncertain of your ability to make friends with him-in each case, fear of failure is almost certain to result in failure. Success, on the other hand, is won most often by those who believe in winning.

Never doubt for a moment that you are going to succeed. Look forward to that success with just as much assurance as you look forward to the dawn of another day, then work-with all that's in you-for success.

139-LT-216

ACHIEVEMENT THROUGH ELECTRONICS

WIDE-BAND AMPLIFIERS

B204

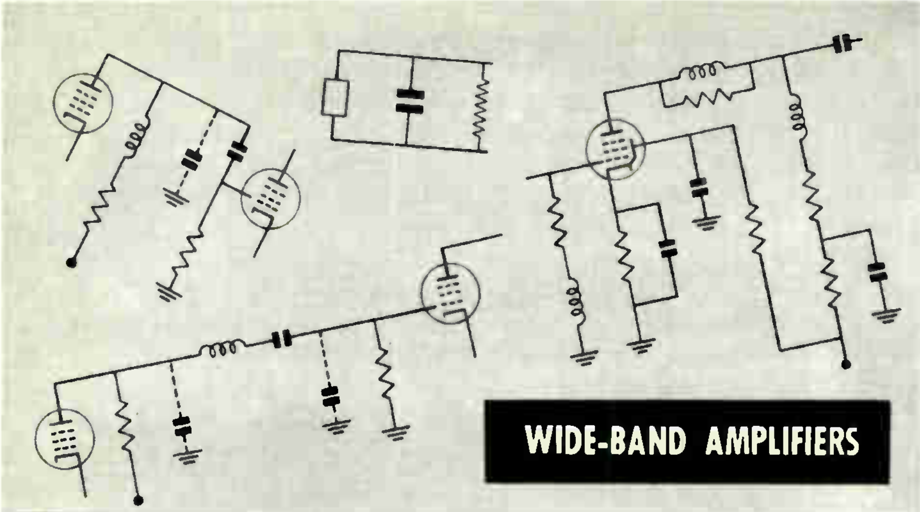
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

WIDE-BAND AMPLIFIERS

B204

STUDY SCHEDULE

KL5M1273 1975 EDITION Litho in U.S.A.



WIDE-BAND AMPLIFIERS

A wide-band amplifier is an amplifier designed to amplify a wide range of signal frequencies. The amplifiers you studied in earlier lessons were designed to amplify only a limited range of frequencies. For example, a low-frequency voltage amplifier such as found in a typical radio receiver or in the sound portion of a television receiver, is designed to amplify only frequencies from about 100 Hertz up to about 10,000 Hertz. Even in a high fidelity amplifier or a high fidelity radio receiver, the low-frequency voltage amplifier is designed to amplify only frequencies from about 10 Hertz up to about 30,000 Hertz. Amplifiers of this type are entirely unsatisfactory for use as video amplifiers in a television receiver. Here, we need a special wide-band amplifier that can amplify all the signals in a TV picture from only a few cycles per second (a few Hertz) all the way up to several megacycles per second (several megahertz). Furthermore, the amplifier must be able to amplify all these signals equally well and must not shift the phase of any signal more than or less than the phase of the other signals.

In your earlier lesson on low-frequency amplifiers you learned that the response of a resistance-capacitance coupled amplifier is limited by three things. The low-frequency response is limited by the reactance of the coupling capacitor. In amplifiers that use vacuum tubes, when the coupling capacitor has a reactance that is equal to or greater than the resistance of the grid resistor in the following stage, the gain of the amplifier begins to drop to such a low level that considerable frequency distortion results. You will remember that we have frequency distortion in an amplifier when signals of different frequencies do not receive the same amount of amplification. In transistor amplifiers we have the same problem when the reactance of the coupling capacitor becomes equal to or greater than the input resistance of the following stage. This input resistance is equal to the input resistance of the transistor itself (which is usually quite low) in parallel with any resistance connected between the base and the emitter.

The high-frequency response is limited by the shunt capacities in the circuit. At high frequencies the shunt capacities in the circuit have a low enough reactance to have an appreciable effect on the gain of the amplifier. In a vacuum tube amplifier, the shunt capacities reduce the effective value of the plate load resistance. In a transistor amplifier they reduce the effective collector load resistance. In both cases the gain of the amplifier drops so that it no longer gives satisfactory results.

Both the low-frequency and the high-frequency response are limited by phase shift. You will remember that the current flowing through a capacitor leads the voltage across it by 90 degrees. Thus, as the capacitive reactance in the circuit becomes appreciable, there is an appreciable phase shift. This means that both low-frequency and high-frequency signals are shifted in phase with reference to middle-frequency signals. In TV receivers and in other electronic devices, this phase shift can appreciably affect the performance of the equipment.

In this lesson we are going to learn what steps are taken to improve both the low-frequency response and the high-frequency response of resistance-coupled amplifiers so that they can be used as wide-band amplifiers.

Another thing we are going to investigate in this lesson is the method used to describe the performance of an amplifier. For example, a manufacturer may say that an amplifier is flat from 100 Hertz to 100 KHz. This means that the gain of the amplifier is constant from 100 Hertz to 100 KHz. In other words, if the voltage gain is 100 at 100 Hertz it will be 100 at all frequencies between 100 Hertz and 100 KHz.

An amplifier that is flat from 100 Hertz to 100 KHz is quite a good amplifier. Usually it is impossible to design an amplifier that has exactly the same gain over a wide range of frequencies. The gain may be a little higher or a little lower at some frequencies than it is at others. The exact variation in gain that can be tolerated depends largely on what the amplifier is to be used for. Therefore, manufacturers must have a method of describing how much the gain of the amplifier varies. To do this they use a unit called the decibel (db). Since the decibel is such a useful unit and since you will encounter it in all branches of electronics, we will learn something about it now before going ahead with our study of wide-band amplifiers. We will then be able to use the decibel in describing amplifier performance so you can see how it is used by manufacturers, engineers, and technicians.

Logarithms and Decibels

The decibel is a logarithmic ratio. In other words, it is a ratio that is based on logarithms. Therefore, before we can understand what a decibel is, we must first learn something about logarithms.

There are two important types of logarithms in use today. One is called a "common" logarithm and the other a "natural" logarithm. Common logarithms are based on the number 10. This is the type of logarithm that we will study and use now.

THE THEORY OF LOGARITHMS

The basic idea of logarithms comes from the fact that any number can be expressed as the "power" of another number. The "power" of a number is the product of a number multiplied by itself a given number of times. The first power of a number is the number itself; the second power is that number multiplied by itself; the third power is that number multiplied by itself twice, etc.

This is easiest to understand by taking an example. In the system of common logarithms, we express all numbers as powers of 10; so we will use 10 as our example. The number 10 itself is equal to 10 to the first power. This can be written 10^1. 100 is equal to 10 X 10. This is 10 to the second power, and can be written 10^2. Similarly, 1000 is equal to 10 X 10 X 10, which is 10 to the third power, and can be written 10^3. The number 10,000 is 10 to the fourth power. Since 10 X 10 X 10 X 10 is equal to 10,000 it can be written 10^4.

Now, as we have said, 100 is 10 to the second power; the logarithm of 100 is simply the power to which 10 must be raised to give us 100. Ten must be raised to the second power (10^2) to give us 100. Therefore, the logarithm of 100 is 2. Similarly the logarithm of 1000 is 3, and the logarithm of 10,000 is 4. The logarithm of 10 is 1.

This is not very complicated when a number is an exact power of 10. But let us consider the numbers between 10 and 100. It is a little more difficult to see how a number between 10 and 100 can be expressed as a power of 10. Actually, this is quite difficult to work out mathematically, but it can be done. Fortunately all these values have been worked out and are available in tables called logarithm or "log" tables. If you want to know the logarithm of a number, you must refer to the table. For example, the logarithm of the number 2 is .301. This means that if it were possible to multiply the number 10 by itself .301 times, the product would be 2. This can be written $10^{.301}$. The exponent, or power, .301, is called the logarithm.

Now let's take the number 20. The logarithm of 20 is 1.301. Notice now that the logarithm is divided into two parts - one part to the left of the decimal point, the other to the right of the decimal point. The part to the left is called the "characteristic" and the part to the right is called the "mantissa".

In the logarithm of a number, the "characteristic," is 1 less than the number of figures to the left of the decimal point in the original number. In other words, any number between 10 and 99 would have a logarithm with a characteristic of 1. The important rule to remember here is that the characteristic is always one number smaller than there are whole numbers in the original number. Thus if the logarithm of a number is exactly three, the number itself is exactly 1000 because 1000 has four places to the left of the decimal.

The chart shown in Fig. 1 gives the characteristics of the numbers you are likely to encounter.

For numbers from:			Characteristic
1	to	9	0.
10	to	99	1. (10^1)
100	to	999	2. (10^2)
1,000	to	9,999	3. (10^3)
10,000	to	99,999	4. (10^4)
100,000	to	999,999	5. (10^5)

Fig. 1. The characteristics of numbers from 1 to 999,999.

A table of logarithms is shown on pages 6 and 7. As you can see, it shows only the mantissas. When you are looking up a logarithm, you must supply the characteristic from the information in Fig. 1. Notice that the table does not list an infinite number of mantissas because logarithms repeat themselves. For example, the mantissa of a logarithm will be the same for the number 2 as it is for the number 20, or 200 or 2000. The only difference in the logarithm will be in the characteristic. For example, the logarithm of the number 2 is .301. The logarithm of 20 is 1.301 and the logarithm of 200 is 2.301. The logarithm table would simply give you a value for the number 2. If the number is 20 or 200 or 2000, you must remember to add the correct characteristic in front of the mantissa. Similarly the mantissa of the logarithm of 21 would be the same as the mantissa of the logarithm of 210, the difference would be in the characteristic. The log of 21 is 1.322 and the log of 210 is 2.322.

Let's take the number 39 and see how we would find the logarithm. We know that the characteristic will be 1, because it is always equal to 1 less than the number of figures in the antilog. (The original number is called the antilog.)

Now, to find the mantissa, refer to the log tables. First find the number (39) in the N column. Since the number 39 is the complete number, follow across to the 0 column. There you will find 5911, which is the mantissa. Since you already know that the characteristic is 1, you have the complete logarithm of 39, that is, 1.5911.

If the number for which you wanted the logarithm had 3 places, you would find the first two in the N column, then follow across to the column under the third digit. For example, to find the logarithm of 399, you would find 39 in the N column, and then follow across to the 9 column, where you would find 6010 for the mantissa. Since the number 399 has 3 places, you know the characteristic is 2, so the complete logarithm is 2.6010.

The last column in the log table is labeled P.P.; this is the "Proportional Parts" column. This column is used when the number has more than three digits. For example, the log of 399 is 2.6010. The log of 3990 has the same mantissa, .6010, but the characteristic is 3, so the log is 3.6010. But what about the log of the number 3995? It is greater than the log of 3990, but less than the log of 4000. The proportional parts column is used to get the log of 3995. We look in this column under the heading 5 and get 6. So to the log of 3990, which is 3.6010 we add .0006 and get 3.6016, which is the log of 3995. The proportional parts column only goes up to 5, so if the fourth digit is greater than 5, we add the proportional parts of two numbers that add up to the fourth digit. For example, to express 3998, we take the log of 3990, which is 3.6010 and add the proportional parts under 3 and 5 (3 and 6) and get 3.6019, which is the log of 3998. We could also take the log of 4000, which is 3.6021 and subtract the value under the 2 column and get 3.6019.

THE DECIBEL

Many years ago engineers working on telephone installations introduced a unit of power measurement called the bel. This unit of measurement was named for Alexander Graham Bell, the inventor of the telephone.

The bel was introduced as a unit of measurement because engineers and scientists discovered that the human ear responds to variations in loudness in an approximately logarithmic manner. Therefore, it is convenient to have a unit that can be used to express the ratio between the power of two signals in a logarithmic manner. The bel is simply the logarithm of the ratio of the power of two signals. For example, if we had a signal power of 100 watts and another signal power of 10 watts, and we wished to express the ratio of these two signals in bels, we would use the formula:

$$\text{bels} = \log \frac{P_1}{P_2}$$

and substitute 100 watts for P_1 and 10 watts for P_2, and we would get:

$$\text{bels} = \log \frac{100}{10}$$

$$\text{bels} = \log 10$$

Now the log of 10 is 1, so this power ratio is equal to 1 bel. In other words, a power ratio of 100 watts to 10 watts, which is a ratio of 10 to 1, is equivalent to 1 bel. Thus a power ratio of 10 watts to 1 watt or 1000 watts to 100 watts are both power ratios of 10 to 1 so they also represent a change in power of 1 bel.

The bel proved to be too large a unit to handle easily, so another unit, one-tenth the size of the bel was introduced. This unit is called the decibel (abbreviated db). Thus the commonly used measuring unit is the decibel; the prefix deci means one-tenth. A power ratio in decibels is defined as:

$$\text{db} = 10 \log \frac{P_1}{P_2}$$

which simply means that the ratio of two powers expressed in decibels is equal to ten times the logarithm of the ratio of the two powers.

You will notice that the above relationship refers to power ratios only. It is common in electronics work to refer to voltage ratios, especially when calculating or discussing the gain of amplifiers. When the ratio between two voltages is calculated in decibels, we must modify the decibel equation to take care of the fact that the power ratios are proportional to the squares of voltage ratios since $P = E^2 \div R$. The formula to express voltage ratios in decibels is:

$$\text{db} = 20 \log \frac{E_1}{E_2}$$

It is important to keep in mind that the voltage formula can be used only when the resistances in the two circuits being compared are equal. If we are trying to compare voltages developed across resis-

N	0	1	2	3	4	5	6	7	8	9	P. P. 1·2·3·4·5
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374	4·8·12·17·21
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755	4·8·11·15·19
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106	3·7·10·14·17
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430	3·6·10·13·16
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732	3·6·9·12·15
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014	3·6·8·11·14
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279	3·5·8·11·13
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529	2·5·7·10·12
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765	2·5·7·9·12
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989	2·4·7·9·11
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201	2·4·6·8·11
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2·4·6·8·10
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2·4·6·8·10
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2·4·5·7·9
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2·4·5·7·9
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2·3·5·7·9
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2·3·5·7·8
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2·3·5·6·8
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2·3·5·6·8
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1·3·4·6·7
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1·3·4·6·7
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1·3·4·6·7
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1·3·4·5·7
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1·3·4·5·6
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1·3·4·5·6
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1·2·4·5·6
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1·2·4·5·6
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1·2·3·5·6
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1·2·3·5·6
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1·2·3·4·6
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1·2·3·4·5
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1·2·3·4·5
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1·2·3·4·5
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1·2·3·4·5
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1·2·3·4·5
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1·2·3·4·5
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1·2·3·4·5
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1·2·3·4·5
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1·2·3·4·4
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1·2·3·4·4
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1·2·3·3·4
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1·2·3·3·4
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1·2·2·3·4
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1·2·2·3·4
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1·2·2·3·4

N	0	1	2	3	4	5	6	7	8	9	P. P. 1·2·3·4·5
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1·2·2·3·4
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1·2·2·3·4
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1·2·2·3·4
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1·1·2·3·4
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1·1·2·3·4
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1·1·2·3·4
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1·1·2·3·4
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1·1·2·3·3
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1·1·2·3·3
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1·1·2·3·3
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1·1·2·3·3
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1·1·2·3·3
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1·1·2·3·3
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1·1·2·3·3
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1·1·2·3·3
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1·1·2·2·3
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1·1·2·2·3
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1·1·2·2·3
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1·1·2·2·3
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1·1·2·2·3
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1·1·2·2·3
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859	1·1·2·2·3
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1·1·2·2·3
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1·1·2·2·3
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1·1·2·2·3
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1·1·2·2·3
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1·1·2·2·3
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1·1·2·2·3
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1·1·2·2·3
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1·1·2·2·3
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1·1·2·2·3
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1·1·2·2·3
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	0·1·1·2·2
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	0·1·1·2·2
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	0·1·1·2·2
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0·1·1·2·2
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0·1·1·2·2
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0·1·1·2·2
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0·1·1·2·2
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0·1·1·2·2
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0·1·1·2·2
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0·1·1·2·2
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0·1·1·2·2
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0·1·1·2·2
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0·1·1·2·2

tors of unequal value, we must convert the voltage to the power developed across the resistors and then use the power formula.

USING THE DECIBEL

It so happens that the smallest amount of change in sound power level that can be distinguished by the average human ear is 1 decibel on a sine wave signal, or 3 db on complex waves such as the average human voice.

Because the decibel is such a convenient unit for expressing changes in sound level, manufacturers of audio equipment have for some time used it in describing the response of their amplifiers. This practice has generally spread into describing the performance of wide-band amplifiers.

Let us now see two examples of how the decibel can be used to describe an amplifier response. In an earlier lesson we discussed the so-called half-power points and the .707 voltage points in an amplifier. Remember that when the reactance of the coupling capacitor used between two resistance-capacitance stages becomes equal to the resistance in the input of the following stage, the voltage gain of the amplifier drops to .707 (70.7%) of what it is at medium frequencies. At the same time, since the voltage drops to .707, the current also drops to .707. The power, which is the product of E X I, is therefore .707 X .707, which is approximately .5 times what it is at the middle frequencies.

Thus we called this point the half-power point. Now let's use decibels to see how much of a change this represents.

Considering first the change in output voltage, let's call E_1, the voltage at the middle frequencies, 1; E_2, the voltage at the low frequency, will then be .707. Thus the change in db is:

$$db = 20 \log \frac{E_1}{E_2}$$

and substituting 1 for $\mathbf{E_1}$ and .707 for $\mathbf{E_2}$ we get:

$$db = 20 \log \frac{1}{.707}$$

Now when we divide .707 into 1 we get the following:

$$.707 \overline{)\ 1\ }$$

We get rid of the decimal in the division by moving the decimal point in each number three places to the right so we have 1000 ÷ 707, which is:

```
          1.414
    707)1000.
         707
         2930
         2828
          1020
           707
           3130
           2828
            302
```

Thus we can see that 1 divided by .707 is equal to 1.414, plus a remainder. For our purposes 1.41 is close enough. Thus we have:

$$db = 20 \log 1.41$$

The log of 1.41 is .1492, which we can call .15, so we have:

$$db = 20 \times .15 = 3$$

Hence the change in db is 3 db. This means that the voltage gain of the amplifier has changed by 3 db. In this case, at the frequency when the reactance of the coupling capacitor is equal to the resistance in the input circuit of the following stage, the voltage output will have dropped 3 db from what it is at the middle-frequency range.

Now let's see what results we get if we use the power formula. Remember that when the voltage drops to .707, the current also drops to .707, and therefore the power drops to .5. Thus, using the formula:

$$db = 10 \log \frac{P_1}{P_2}$$

and substituting 1 for P_1 and .5 for P_2 we get:

$$db = 10 \log \frac{1}{.5}$$

and since $1 \div .5 = 2$, we have:

$$db = 10 \log 2$$

The log of 2 is .301, and 10 times this is 3.01, or for all practical purposes, 3 db. Thus, whether we use the power formula or the voltage formula we get the same change in db.

Manufacturers frequently use decibels to express the change in power output over a given frequency range in a power amplifier or the change in voltage output in a voltage amplifier. For example, in describing a voltage amplifier, the manufacturer might say that the voltage gain of the amplifier is flat within 3 db from 10 Hertz to 3 megahertz. This means that the voltage gain of the amplifier does not vary by more than 3 db above or below the middle-frequency gain of the amplifier between the frequencies of 10 Hertz and 3 MHz. You know that 3 db represents a voltage change of .707, or in other words, the gain of the amplifier will not vary more than 29.3% (100% - 70.7%) from what it is at the middle frequencies. The gain will be at least 70.7% of the middle-frequency gain within the frequency range of from 10 Hz to 3 MHz.

The manufacturer of a certain power amplifier might claim that the power output from the amplifier is flat within a certain number of db from 50 Hertz to 1000 Hertz. This means that the power output between these frequency limits is within the specified number of db of the specified power output.

One of the advantages of using the decibel in comparing power ratios is that it gives a pretty good picture of how the amplifier will sound. For example, if you have a 5-watt amplifier that is capable of putting out 5 watts of audio power in the middle-frequency range, but only 1 watt at a frequency of 100 Hertz, you would have a power ratio of 5 to 1. This represents a change of 7 db, which would be very noticeable, even though the actual difference in power output is only 4 watts. On the other hand, if you had an amplifier capable of putting out 100 watts of audio power at the middle frequencies, and dropped to 50 watts at 100 Hz, although the change in power is actually 50 watts, the power ratio is 2 and the db change only 3 db. This is a smaller change in db than the change from 5 watts to 1 watt. This is as it should be, because you would notice a greater change in going from 5 watts to 1

Voltage Ratio	DB	Power Ratio	DB
1	0	1	0
2	6.0	2	3.0
3	9.6	3	4.8
4	12.0	4	6.0
5	14.0	5	7.0
6	15.6	6	7.8
7	16.8	7	8.4
8	18.0	8	9.0
9	19.2	9	9.6
10	20.0	10	10.0
20	26.0	20	13.0
30	29.6	30	14.8
40	32.0	40	16.0
50	34.0	50	17.0
60	35.6	60	17.8
70	36.8	70	18.4
80	38.0	80	19.0
90	39.0	90	19.6
100	40.0	100	20.0
1,000	60.0	1,000	30.0
10,000	80.0	10,000	40.0

Fig. 2. Decibel values corresponding to voltage and power ratios.

watt than you would in going from 100 watts to 50 watts.

The chart shown in Fig. 2 is a table of decibel values corresponding to voltage and power gains. This would give you an idea of what the voltage ratio or the power ratio is for certain db values.

The decibel is an important unit in the electronic field. You should be familiar with what it is and how it is used. If you plan on doing radio and TV service work you do not have to be able to calculate either voltage gain or power ratios in decibels. However, you should realize what the decibel is, and become familiar with its use. As a technician you will run into it time and time again. Manufacturers frequently use it in describing the performance of electronic equipment. You'll be called on to interpret the meaning of characteristics of this type. Also an understanding of what a decibel is and how it is used will help you in evaluating the performance of certain types of electronic equipment. By comparing an amplifier with the manufacturer's specifications, you can decide whether or not the amplifier is performing as well as it is supposed to be able to.

If you intend to go into communications or into industry as an electronics technician you should be able to calculate both voltage gain and power ratios in decibels. The Self-Test Questions that follow will give you an opportunity to try to perform these calculations; you can check your answers with those given.

SELF-TEST QUESTIONS

(a) What is a common logarithm?

(b) What is the part of the logarithm to the left of the decimal point called?

(c) What is the part of the logarithm to the right of the decimal point called?

(d) What is the characteristic of numbers from 100 to 999?

(e) Write the logarithm of 7.

(f) Write the logarithm of 700.

(g) Write the logarithm of 41.7.

(h) What is the bel?

(i) What is the decibel?

(j) If the power output of an amplifier changes from 222 watts to 37 watts, how many db does this change represent?

(k) If a defect develops in an amplifier and the output voltage drops from 150 volts to 75 volts, what change in decibels does this represent?

Extending the High-Frequency Response of an Amplifier

A schematic diagram of a typical two-stage R-C coupled amplifier using vacuum tubes is shown in Fig. 3. The first stage, V_1, uses a triode tube, and the second stage, V_2, uses a pentode tube. In your study of low-frequency amplifiers, you learned that the high-frequency response of an R-C coupled amplifier is limited by shunt capacities in the circuit. You will remember that the first tube, V_1, has a certain capacity between the plate and cathode of the tube. This capacity in effect is in parallel with the plate load resistor R_3. You will remember that as far as the signal is concerned, R_3 is connected between the plate of the tube and ground. The end of R_3 that connects to the positive side of the power supply is at signal ground potential because the output filter capacitor in the power supply will be so large that at signal frequencies it acts like a short circuit. The cathode of the tube is adequately bypassed by the bypass capacitor C_1 so that the plate-to-cathode capacity of V_1 is in effect connected directly between the plate of the tube and ground, or in parallel with R_3.

There is also another capacity that must be considered in the input circuit of V_2. This capacity is the grid-to-cathode capacity of the tube. This capacity will depend upon the tube structure and also upon the gain of the stage. At high frequencies, the coupling capacitor C_2 has such a low reactance that it can be ignored so that in effect the output capacity of V_1 is connected directly in parallel with the input capacity of V_2.

In addition to the capacities in the tubes V_1 and V_2, there will be a certain amount of stray capacity in the wiring. For example, the coupling capacitor C_2 will have a certain capacity to ground, the lead connecting C_2 to V_1 and V_2 will also have a certain capacity to ground. All this capacity is added to the capacity in the output circuit of V_1 and the capacity

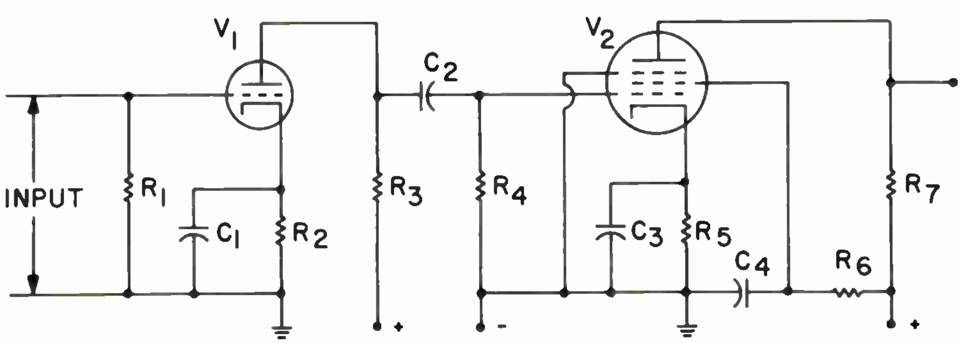


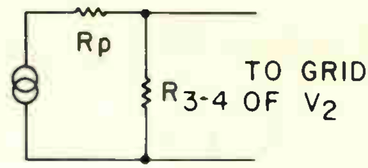
Fig. 3. An R-C coupled amplifier.

in the input circuit of V_2, and tends to limit the high-frequency response of the amplifier.

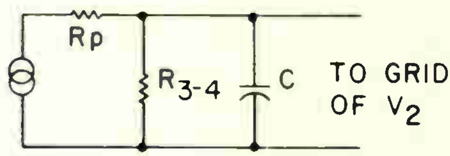
EFFECT OF SHUNT CAPACITY

In Fig. 4 we have shown the equivalent circuits of the coupling network used between V_1 and V_2 at middle frequencies and at high frequencies. The circuit in A is the equivalent circuit of the coupling network at middle frequencies. The resistor $R_{3\text{-}4}$ represents the resistance of R_3 and R_4 in parallel. As far as the signal is concerned, they are in parallel because the reactance of C_2 is so low that the capacitor acts like a short circuit. The parallel combination of R_3 and R_4 will be almost equal to the resistance of R_3 because R_4, which is the grid resistor of V_2, is many times larger than the plate resistor R_3. Therefore as far as the load in the plate circuit of V_1 is concerned, it is almost entirely controlled by the value of R_3. R_3 will usually have a resistance somewhere between 10,000 and 100,000 ohms. R_4, on the other hand, will be considerably larger than this, usually between .25 and a .5 megohm. The value of R_4 is limited by the type of tube used for V_2. Some tubes will develop a negative bias due to electrons accidentally striking the grid if too large a resistance is used in the grid circuit. Other tubes may have a small amount of gas present inside of the tube; these tubes will develop a positive voltage on the grid if too large a grid resistor is used. However, in any case, even though R_4 is in parallel with R_3 at middle audio frequencies, it has very little effect insofar as reducing the plate load is concerned because its resistance is usually considerably greater than that of R_3.

In Fig. 4B we have shown the equivalent circuit of the coupling network at high frequencies. The capacity C is used to represent the sum of the output



(A)



(B)

Fig. 4. Equivalent circuits of coupling networks between V_1 and V_2. "A" shows the equivalent circuit at middle frequencies and "B" the equivalent circuit at high frequencies.

capacity of V_1, the input capacity of V_2, and the distributed capacities in the circuit. Notice that C is connected directly across $R_{3\text{-}4}$.

As the frequency of the signal amplified by the amplifier increases, the capacitive reactance of C decreases. You will remember that the capacitive reactance of the capacitor is given by the formula:

$$X_C = \frac{1}{6.28 \times f \times C}$$

From the formula you can see that if either f or C increases, the value of the capacitive reactance will decrease.

At some frequency the capacitive reactance of C will be equal to the resistance of the parallel combination of $R_{3\text{-}4}$. You will remember that we pointed out previously that when this happens the signal reaching the grid of V_2 will actually be 70.7% of the signal reaching the grid of V_2 at middle audio frequencies. You will remember we called this the half-power point; this is the point at which the signal reaching the grid of V_2 is 3 db lower than the signal reaching the grid of V_2 at middle audio frequencies. Now let us see the steps we can take to increase the frequency at which the signal reaching the grid of V_2 drops 3 db.

REDUCING THE SHUNTING EFFECT

In a practical amplifier that is to be designed to amplify signals of widely different frequencies, we try to keep the response of the amplifier as flat as possible over the required frequency range. This means that we try to design the amplifier so that it will give all the signal frequencies it must amplify the same amount of amplification. If all the signal frequencies receive the same amount of amplification, we say that the amplifier is flat over the range of frequencies it must amplify. However, in actual practice it is impossible to get an amplifier that is exactly flat over a very wide frequency range, so we generally have to be satisfied with an amplifier whose gain is flat within a few db over the frequency range for which the amplifier is designed.

To give a practical example of how the high-frequency response of an amplifier can be extended, let us see what we can do to extend the frequency response of an amplifier like the one shown in Fig. 3. Suppose that we want to be able to amplify signals up to 200 KHz so that the gain of the amplifier at 200 KHz is within 3 db of the gain at the middle frequency. If in checking the amplifier we find that the gain drops 3 db at 100 KHz, we can extend the frequency response of the amplifier by reducing the size of the plate load resistor R_3.

If the gain drops 3 db at 100 KHz, this means that the capacitive reactance of the shunt capacity across the parallel combination of R_3 and R_4 must be equal to the value of the parallel resistors at a frequency of 100 KHz. You will remember that R_4 is much larger than R_3. Therefore, the value of the combination will be almost equal to the value of R_3 alone. If we cut the value of R_3 in half, we will in effect be cutting the value of R_3 and R_4 in parallel, in half. Then the reactance of the capacitance will be equal to $R_{3\text{-}4}$ at twice the original frequency of 100 KHz. Therefore by simply reducing the plate resistor R_3 in the plate circuit of V_1, we can extend the gain of the amplifier at high frequencies. Of course, what we are actually doing is reducing the gain of the amplifier at the low and middle frequencies in order to flatten the response of the amplifier over a wider frequency range. In spite of the fact that this might seem to be somewhat of a disadvantage, it is the method most widely used to extend the high-frequency response of an amplifier. If in reducing the amplifier gain by reducing the plate load resistance, we find that we do not have sufficient gain in the amplifier, then we can overcome this difficulty simply by adding an additional stage.

In video amplifiers found in television receivers, where the frequency response must be reasonably flat up to a frequency of several megahertz, the plate load resistor R_3 may have a resistance of only a few thousand ohms. This makes it possible to obtain a reasonably flat gain over a wide frequency range. Suppose for example, that the value of R_3 is 94,000 ohms and that the gain of the amplifier drops 3 db at 100 KHz. By reducing the size of R_3 to 47,000 ohms, we can extend the 3 db down point to 200 KHz. If we reduce the value of R_3 to 4,700 ohms, then we can extend the 3 db down point to 2 MHz. Of course, the gain of the amplifier will be much lower at low and middle frequencies with a 4700-ohm resistor in the plate circuit of V_1 than it will be with a 47,000-ohm or a 94,000-ohm resistor in the plate circuit. But the gain will be essentially flat from a very low frequency up to approximately 2 MHz with the 4.7K resistor in the plate circuit. Where a flat response over a wide frequency range is required, reducing the size of the plate load resistor in order to reduce the shunting effect of the shunt capacity is the simplest way of obtaining the wide frequency response. When we use a very small value of plate load resistor, it is so small compared to the resistance of the grid resistor in the following stage, that we can forget the shunting effect of the grid resistor and consider the plate load as equal to the value of the plate load resistor. In the next section dealing with high-frequency compensation we will do this.

In the example we have given, we found that by reducing the size of R_3 to 4700 ohms we can extend the 3 db down point to 2 MHz. However, in some applications we may want to keep the gain essentially flat out to 2 MHz or higher. We cannot reduce the value of R_3 much below 4700 ohms or we'll get very little gain from V_1. Therefore, we have to use another method of extending the high-frequency response of the amplifier. We can do this by high-frequency compensation.

HIGH-FREQUENCY COMPENSATION

A method widely used to improve the response of an amplifier at high frequencies is called compensation. By compensation we mean that we add something to the circuit to compensate for other undesirable effects. As you might guess, since the capacity and capacitive reactance are the causes of difficulty at high frequencies, to counteract this, we add inductance that will introduce inductive reactance into the circuit.

The coils that are added to the circuit to improve the high-frequency response are called peaking coils. They are so-called because they will peak the response at some frequency above the maximum frequency that could be amplified by the amplifier without these coils. There are actually three types of circuits that can be used. There is a circuit known as shunt peaking, one known as series peaking and

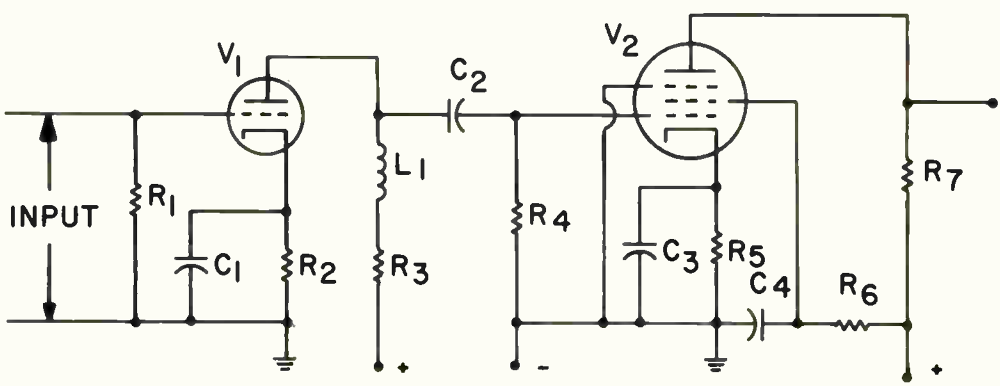


Fig. 5. A wide-band amplifier with shunt-peaking coil.

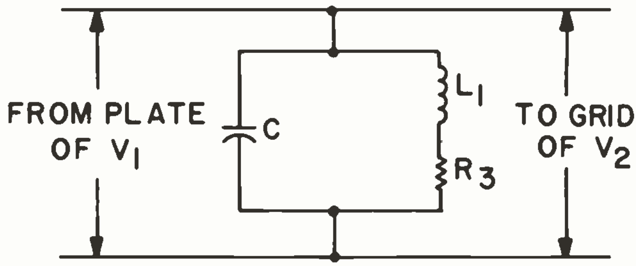


Fig. 6. Equivalent circuit of coupling network between V_1 and V_2.

one that is a combination of shunt and series peaking. The combination of shunt and series peaking is most widely used, but because shunt peaking and series peaking alone are somewhat simpler to understand than the combination, we will look at these two types of peaking-circuits first.

Shunt Peaking. Fig. 5 is a schematic diagram showing how shunt peaking can be added to the R-C coupled amplifier circuit shown in Fig. 3. Notice that the two circuits are identical except that L_1 has been added in series with R_3 in the plate circuit of V_1. The equivalent circuit of the coupling network between V_1 and V_2 with this peaking coil added is shown in Fig. 6.

Notice that in Fig. 6 we see that the peaking coil and load resistor are in parallel with the capacity C which represents the output capacity of V_1, the input capacity of V_2, and the distributed capacity in the circuit. The idea in back of the peaking coil is to select a value of inductance so that a parallel-resonant circuit is formed at a frequency above the frequency at which the gain would drop 3 db without the peaking coil in the circuit. Then, as the frequency of the signal to be amplified approaches the 3 db down point, without the peaking coil, the parallel-resonant circuit begins to take over. You will remember that one of the characteristics of the parallel-resonant circuit is that it acts like a high resistance at resonance. Therefore instead of the load impedance of V_1 dropping because of the shunting effect of the capacity, the load impedance actually begins to increase because of the high resistance of the parallel-resonant circuit. This will cause the gain of the amplifier to increase slightly so that higher frequency signals can be amplified with the same gain as signals at middle frequencies.

At first you might think that the resistance R_3 that is in series with the peaking coil, would lower the Q of the parallel-resonant circuit so that the circuit would not be particularly affected. The resistance does in fact lower the Q of the resonant circuit, but this is desirable. The purpose of L_1 is to keep the gain of the amplifier flat at higher frequencies. R_3 reduces the resistance of the parallel-resonant circuit and tends to keep the load in the plate circuit of V_1 more or less constant. If R_3 does not lower the Q of the resonant circuit sufficiently, then we have a situation where the gain rises as the combination of L_1 and C approach resonance. We'll actually have a peak in the response at resonance, if the coil is not loaded sufficiently, so that the gain in the circuit is much higher at this fre-

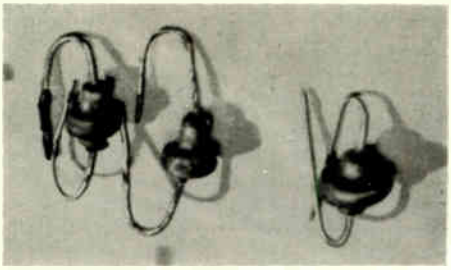


Fig. 7. Peaking coils are wound on ½ watt resistors.

quency than it is at lower frequencies. This is called over compensation, and in most cases it is undesirable. A small amount of over compensation, however, is sometimes used in order to produce some desirable effects. For example, a small amount of over compensation in the video amplifier of a television receiver may tend to make the fine detail in the picture somewhat sharper. However, extensive over compensation will cause ringing where oscillation will occur in the resonant circuit. In a television picture, this would result in fine details being repeated. In other words, if there was a vertical pole appearing in a certain scene, a second or third pole might appear in the picture displaced slightly to the right of the original pole.

Peaking coils are often wound on 1/2 watt resistors If there is not sufficient loading of the resonant circuit to prevent ringing or over peaking, a comparatively low-resistance resistor is used. This provides additional loading on a resonant circuit. Where the circuit already has sufficient loading, the coil can be wound on a dummy form which looks like a 1/2 watt resistor but actually has no electrical connection through it. In some cases, rather than go to the trouble of getting dummy forms on which to wind a coil of this type, manufacturers will simply wind the coil on a very high value resistance. The resistance of the resistor is so high that it has no appreciable loading effect on the circuit. A number of typical peaking coils are shown in Fig. 7.

Series Peaking. Two circuits using series peaking are shown in Fig. 8. The series-peaking coil is labeled L_1 in both circuits. Although L_1 is placed in a slightly different position in the two circuits, the net electrical effect is the same. Coil L_1 isolates the output capacity of V_1 from the input capacity of V_2. At the same time, the value of L_1 is selected so that it will resonate with the input capacity of V_2 at a frequency near or slightly above the frequency at which the gain of the amplifier would drop to 70.7% of the mid-frequency gain without compensation.

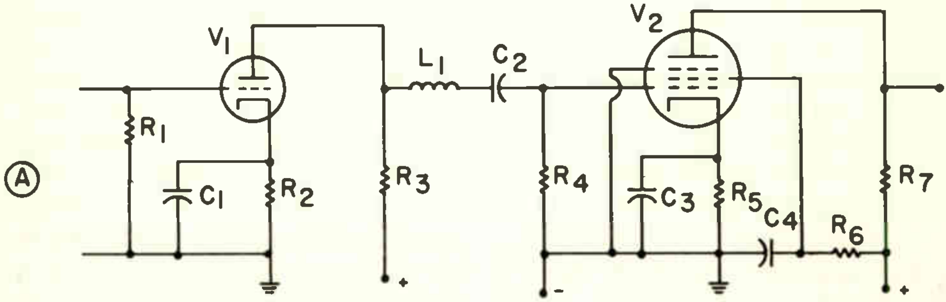


Fig. 8A. Series-peaked wide-band amplifier.

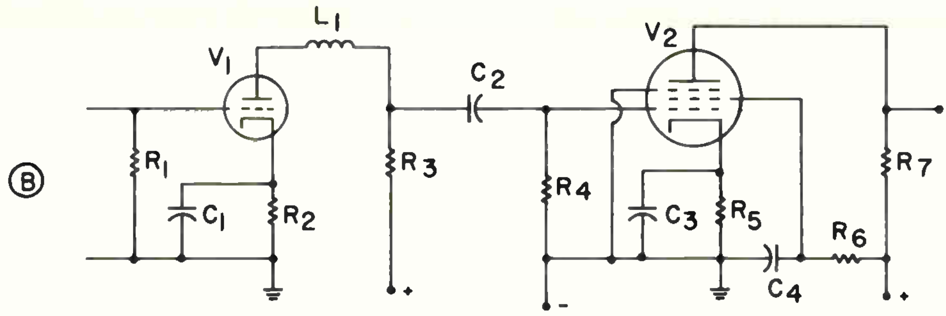
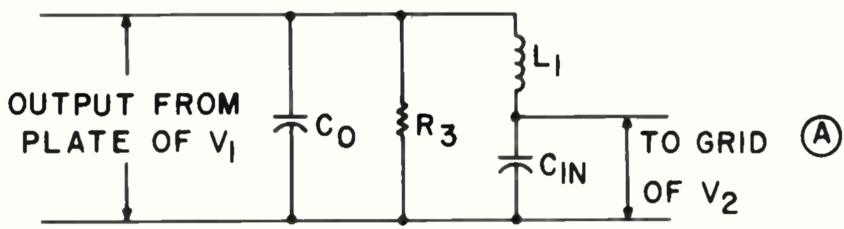


Fig. 8B. Series-peaked wide-band amplifier similar to that shown in Fig. 8A.

In Fig. 9 we have shown the equivalent circuit of the coupling network. Fig. 9A shows the equivalent of the circuit shown in Fig. 8A. Notice that here we have the capacitor labeled C_O. This represents the output capacity of V_1. This capacity is in parallel with the plate load resistor R_3. Since the output capacity of V_1 represents only a fraction of the total capacity made up of the output capacity, plus the input capacity of V_2, plus the wiring capacity in the amplifier, the gain of the amplifier will not fall off so rapidly. In other words, we have effectively reduced the capacity that is shunting R_3. At the same time, by connecting L_1 into the



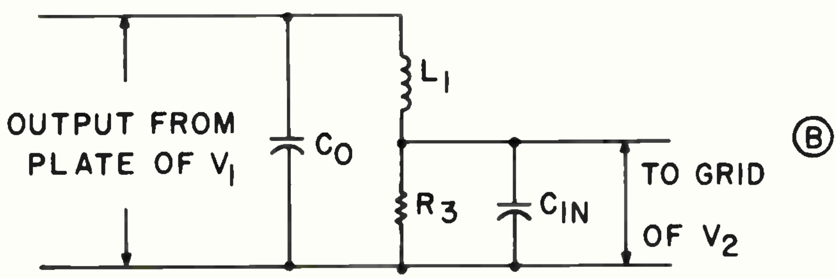


Fig. 9. Equivalent circuits of coupling network in Fig. 8.

circuit (selecting it so that we will resonate it with the input capacity of V_2), we have formed a series-resonant circuit consisting of L_1 and the input capacity of V_2. You know that in a series-resonant circuit there will be a resonant voltage step-up at the resonant frequency. Therefore we will have a resonant voltage step-up across C_{in} at the resonant frequency. This means that although there may be some tendency for the voltage across R_3 to drop at this frequency (because of the shunting effect of the output capacity of V_1), the resonant circuit is able to more than compensate for this drop-off in output from V_1, so that the input voltage of V_2 may actually be higher at high frequencies than the input voltage to V_2 at middle frequencies.

It is usually possible to obtain better peaking with a series-peaking coil than with a shunt-peaking coil. The input capacity of V_2 is usually much larger than the output capacity of V_1, so that splitting the two capacities by means of the peaking coil results in increased output from V_1. At the same time, by using the coil with a high Q, a high resonant voltage step-up can be obtained so that the input signal to V_2 will be substantially boosted. The resistor R_3, which is across the resonant circuit, loads the circuit to prevent over compensation. In some circuits the peaking coil may be loaded by winding it on a resistor to reduce the Q of the circuit still further when the value of the plate load resistor is too high to prevent excessive over compensation. As you will remember, over compensation results in excessive increase in amplification at the high frequency to which the circuit is peaked, and in most cases this is to be avoided.

In Fig. 9B we have shown the equivalent circuit for the circuit shown in Fig. 8B. Here the circuit is somewhat different from the circuit shown at A. In this circuit the load resistor R_3 is across C_{in} at the high frequencies instead of across the entire series circuit. Again at high frequencies the peaking coil L_1 forms a series resonant circuit with the input capacity in the grid circuit of V_2. The resonant voltage step-up in the series circuit tends to compensate for the drop in gain from V_1 due to the reduced size of the plate load.

In servicing amplifiers where peaking coils are used, sometimes you may suspect that a peaking coil is open. If you check across the peaking coil with an ohmmeter, you should get a resistance reading of only a few ohms. If you get a resistance reading of several thousand ohms it indicates that the coil is open and that you are reading through the resistor on which the peaking coil is wound. Usually when a peaking coil is open, it is open right at the end of the coil where the coil is connected to the resistor lead. If you can find the place where the coil connects to the resistor lead, resolder it and if the connection is poor this should clear up the trouble. If on the other hand, it doesn't clear up the trouble, sometimes you can find the wire going from the resistor lead to the coil and see where it is broken. If there isn't enough wire to stretch over to make a connection to the resistor lead, unwinding one turn of the coil will usually enable you to make the connection. Taking a single turn off the coil will not affect its inductance enough to upset its performance in the circuit.

Shunt-Series Peaking. The most satisfactory peaking arrangement is a combination of both shunt and series peaking. You'll find this type of peaking is widely used in the video amplifiers of television

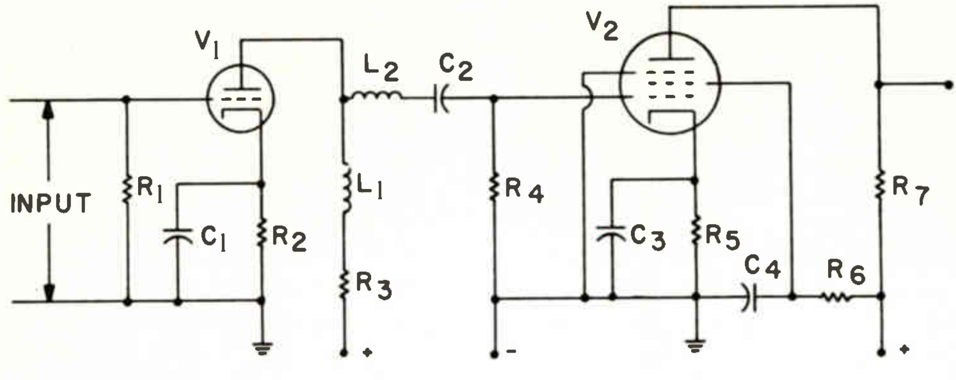


Fig. 10. A shunt-series compensated amplifier.

receivers and in other amplifiers where a wide frequency response is required. The schematic diagram of an amplifier using series-shunt peaking is shown in Fig. 10. This is the same basic amplifier circuit that we started with in Fig. 3, but the shunt-peaking coil, L_1, and the series-peaking coil, L_2, have been added.

An equivalent circuit of the coupling network is shown in Fig. 11. Notice that in the previous case where we used shunt peaking alone, the coil L_1 forms a parallel resonant circuit. However, this time instead of forming a parallel-resonant circuit with the entire capacity in the circuit, it forms a parallel-resonant circuit with the output capacity of V_1. The peaking coil L_2 effectively separates the output capacity of V_1 from the input capacity of V_2. L_2 forms a series-resonant circuit with the input capacity of V_2.

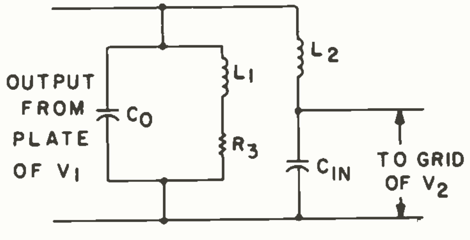


Fig. 11. Equivalent circuit of coupling network in Fig. 10.

By using the series-peaking coil L_2 in this way, we have less capacity in the plate circuit of V_1. This often enables us to increase the size of R_3 which will increase the gain at low and middle frequencies. At the same time, by using the shunt-peaking coil L_1, we can increase the gain of the amplifier at the higher frequencies where it would normally start to fall off. Thus we can maintain a constant output voltage from V_1 over a wider frequency range.

The series-peaking coil, L_2, in addition to separating the output capacity of V_1 from the input capacity of V_2, also forms a series-resonant circuit with the input capacity of V_2. Therefore, as the output signal starts to fall off from V_1, the resonant circuit made up of L_2 and the input capacity of V_2 takes over, and by means of the resonant voltage step-up can keep the signal fed to the grid of V_2 essentially constant.

In circuits using both series and shunt peaking we can select the values of the

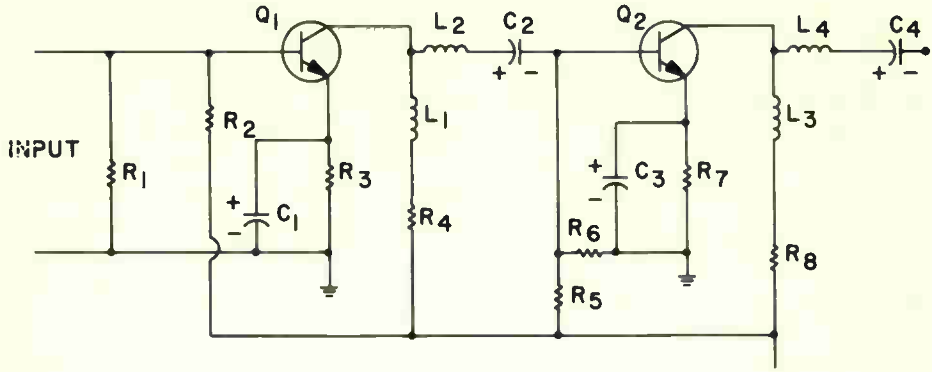


Fig. 12. A series-shunt compensated transistor amplifier.

peaking coils so the resonant circuits resonate at slightly different frequencies. Thus it is possible to make one resonant circuit take over (usually the parallel-resonant circuit) when the gain of the amplifier first starts to fall off. At a frequency above the resonant frequency of the parallel-resonant, the series circuit becomes resonant and compensates for the drop in impedance in the plate circuit of V_1 above the resonant frequency of the shunt-peaking circuit. Thus by using a combination of shunt and series peaking it is usually possible to get more gain from the amplifier over this entire bandwidth because we can use a larger value of load resistor, and also it is easier to extend the high-frequency response of the amplifier and maintain the response essentially flat over a wider frequency range.

TRANSISTOR AMPLIFIERS

So far in our discussion we have been talking about extending the high-frequency response of tube amplifiers. Essentially the same problems exist in transistor amplifiers as exist in tube amplifiers. The output capacity of one transistor, the input capacity of the second transistor, and the stray capacity in the coupling network coupling the two transistors together, limits the high frequency response of the amplifier.

A typical two-stage series-shunt compensated transistor amplifier is shown in Fig. 12. Notice that we have a shunt-peaking coil, L_1, in the collector circuit of Q_1. We have a series-peaking coil connected between the collector of Q_1 and the coupling capacitor C_2. The equivalent circuit of the coupling network used between Q_1 and Q_2 at high frequencies, where the capacity of C_2 can be ignored, is shown in Fig. 13. Notice that we have essentially the same circuit

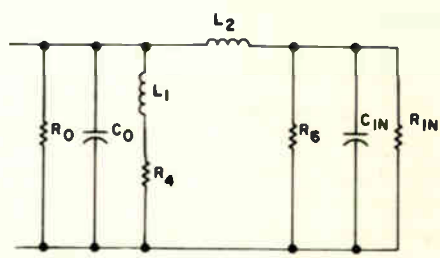


Fig. 13. The high-frequency equivalent of the coupling network used in Fig. 12.

as we had in the series-shunt compensated vacuum-tube coupling circuit except for the resistors R_o and R_{in}. R_o represents the output resistance of Q_1 and R_{in}, the input resistance of Q_2. In the tube circuit, these resistances were so high they could be ignored. However, in transistors, these resistances are low and must be considered.

The peaking coil L_2 separates the output capacity of Q_1 and the associated capacity of the circuit from the input capacity of Q_2 and its associated stray capacity. L_2 is selected to resonate with the capacity in the input circuit of Q_2 and form a series-resonant circuit in order to peak the input to Q_2 at the frequency desired. At the same time, L_1 forms a parallel-resonant circuit in the collector circuit of Q_1 and helps keep the output from Q_1 up at higher frequencies.

We mentioned earlier that the collector load resistor, R_4, is in parallel with the input resistance of Q_2. You will remember from your study of low-frequency voltage amplifiers, that in a transistor circuit Q_2 has a certain base current. This lowers the effective input resistance of the second transistor amplifier, and since this is in effect in parallel with R_4, it has the effect of lowering the value of the collector load resistor in the collector circuit of Q_1. Therefore we already have a comparatively low collector load resistor and there is a limit to how far we can reduce this resistor in order to level out the response of the coupling network over a wide frequency range. You will remember that in the vacuum-tube coupling network we were able to drop the value of the plate load resistor in the first stage and in so doing bring down the low and middle-frequency gain of the amplifier. At the same time, we extended the frequency at which the gain of the amplifier begins to drop off appreciably. Since the collector load resistor is in effect already low, we cannot lower it much further. Fortunately, since the load resistor is low, it takes a higher shunt capacity to have an appreciable effect at high frequencies. Therefore the transistor coupling network tends to give equal response at a higher frequency even without the peaking coils in the circuit than in the case of a vacuum-tube network.

The output of Q_2 may be connected to another amplifier stage, or in a television receiver it might be connected to the picture tube. Additional peaking will be used in the output circuit of this stage in order to keep up the high-frequency response in this circuit. L_3 is a shunt-peaking coil, and L_4 a series-peaking coil as in the network between Q_1 and Q_2.

We run into an additional problem in transistor amplifiers at high frequencies that we do not encounter in vacuum-tube amplifiers. In a common-emitter circuit such as used for Q_1 and Q_2 in Fig. 12, we normally have a 180° phase shift in each stage. You will remember that when the input signal fed to Q_1 drives the base in a positive direction, it increases the forward bias across the emitter-base junction of the NPN transistor. This causes the number of electrons crossing the emitter-base junction to increase and hence the number of electrons reaching the collector will increase. When the number of electrons reaching the collector and flowing through L_1 and R_4 increases, this will cause the voltage drop across R_4 to increase. Therefore the voltage at the collector will swing in a negative direction. Thus with the input signal as shown in Fig. 14A, the collector voltage swings in the opposite direction as shown in Fig. 14B. We say the two signals are 180° out-of-phase.

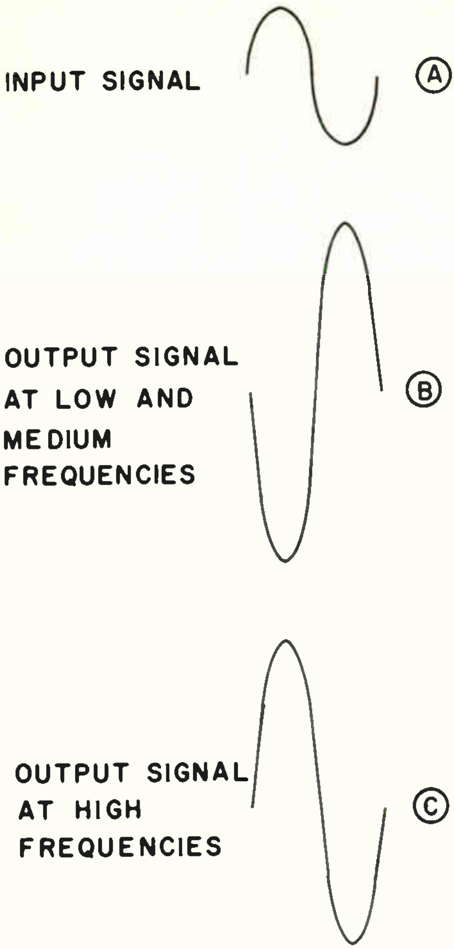


Fig. 14. Phase relationships between input and output signals at low, medium and high frequencies in common emitter amplifiers.

The base of a germanium transistor is usually about one thousandth of an inch thick. The base of a silicon transistor is about one ten-thousandth of an inch thick. It takes the electrons a certain length of time to cross the base of the transistor. It will take the electrons longer to cross the base of the germanium transistor because it is thicker than the silicon transistor. Therefore consider what can happen when the time it takes the electrons to travel through the base is considered. If the signal applied to the base is a low-frequency signal or a medium-frequency signal, the time it takes the electrons to cross the base compared to the time of one cycle is relatively short. The output signal will be 180° out-of-phase with the input signal as shown in A and B of Fig. 14. However, at high frequencies, the electrons may be delayed sufficiently in travelling through the base of the transistor to cause a phase shift between the output signal voltage at medium and low frequencies and the output signal voltage at high frequencies. As a matter of fact, at some high frequency it will take the electrons so long to cross the base that instead of the signals being 180° out-of-phase, they will be in-phase because the electrons are delayed by one half cycle in travelling through the base region. When this happens, the output voltage in the collector circuit of Q_1 will be in-phase with the input voltage applied to the base of Q_1.

You will remember that in a transistor there is a capacity between the collector and the base. A signal is fed from the collector of the transistor back to the base through this capacity. When the output signal is 180° out-of-phase with the input signal, the signal fed back to the base through the collector-base capacity, simply reduces the amplitude of the input signal. This is a form of degeneration; the output from the transistor would not be as high as it would be without this feedback signal. However, at high frequencies where the output signal may be in-phase with the input signal, the signal fed from the collector back to the base will be in-phase with the input signal. This will reinforce the input signal which in turn produces a higher amplitude output signal which in turn builds up the input signal still further. Thus the transistor may go into oscillation and the output signal fed back to the input circuit will take control of the transistor so that it begins producing a high-frequency signal without any input. In circuits where there is a possibility of oscillation occurring, some kind of neutralization must be employed so that a signal that is 180° out-of-phase with the feedback signal, can be fed back into the base circuit to cancel out the signal fed from the collector back to the base through the collector-base capacitance.

FIELD-EFFECT TRANSISTORS

Both the junction-type and the insulated-gate field-effect transistors can be used in wide-band amplifiers. The problems in an amplifier using field-effect transistors are almost identical to those encountered in equipment using vacuum tubes. The output capacity of the first stage plus the input capacity of the second stage along with stray wiring capacity tend to limit the high-frequency response of the amplifier.

A typical two-stage compensated amplifier using junction-type field-effect transistors is shown in Fig. 15. Notice the similarity between this and a vacuum-tube two-stage amplifier.

The peaking coil L_1 in the drain circuit of Q_1 is used to form a parallel-resonant circuit in the drain circuit, to keep the output from Q_1 from falling off at high frequencies. The series-peaking coil L_2 is used to resonate with the input capacity of Q_2 to form a series-resonant circuit to keep up the amplitude of the high-frequency signals fed to the gate of Q_2. In the output circuit of Q_2 we have shown series peaking used between the drain and the following stage.

In many amplifiers using field-effect transistors you will find that series peaking only is used. The one type of peaking is frequently all that is required to obtain the high frequency response required. The value of the resistor R_3, in the drain circuit of Q_1, may be reduced in order to

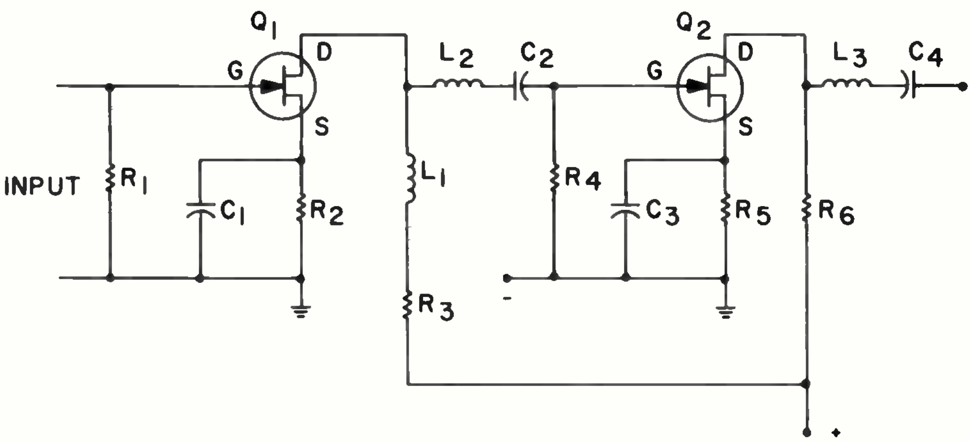


Fig. 15. Compensated amplifier using field-effect transistors.

increase the frequency at which the capacity shunting the output of Q_1 begins to have an appreciable effect on the impedance of the drain circuit of Q_1. You have essentially the same situation as you had with the resistor in the plate circuit of a tube-type amplifier. Reducing the size of the resistor increases the frequency at which the reactance of the output capacitance becomes equal to the resistance of the resistor.

One of the big advantages of the field-effect transistor is that it has a high input resistance. Therefore, the input resistance of Q_2 has little or no loading effect on the output resistance of Q_1. In fact, as far as the performance of the circuit is concerned, it is practically identical to the performance of a two-stage vacuum tube amplifier. Field-effect transistors are able to combine most of the advantages of vacuum tubes along with most of the advantages of transistors and should become increasingly important in the future in electronic equipment.

SELF-TEST QUESTIONS

(l) Why is the value of the plate-load resistor in a wide-band amplifier kept low?

(m) Why is there a limit to the size of grid resistor that can be used in a vacuum-tube amplifier?

(n) How does L_1 in the circuit shown in Fig. 5 help improve the high-frequency response of the amplifier?

(o) How does the peaking coil L_1 in the circuit shown in Fig. 8 help improve the high-frequency response of the amplifier?

(p) What do we mean by series-shunt peaking?

(q) Does capacity shunting have as great an effect on high-frequency response in transistor amplifiers as it does in vacuum-tube amplifiers?

(r) Are there any problems encountered at high frequencies in transistor amplifiers that are not likely to be encountered in vacuum-tube amplifiers?

(s) Are the high-frequency problems encountered in field-effect transistors more like those encountered in vacuum tubes or like those encountered in bipolar transistors?

Extending the Low-Frequency Response of an Amplifier

You will remember that when we studied low-frequency amplifiers we pointed out that the output from an amplifier falls off at low frequencies as well as at high frequencies. The low-frequency response of an amplifier falls off for several reasons. One reason is the reactance of the coupling capacitor used to couple the two stages together increases as the frequency decreases. At some low frequency, the reactance of the capacitor will become so high that it will act like a voltage divider along with the input resistance in the second stage of the amplifier. Thus not all of the amplified signal produced by the first stage is fed to the input of the second stage.

In Fig. 16 we have shown schematic diagrams of three R-C coupled amplifiers. The circuit shown in A is made up of two vacuum tubes, the one at B has two transistors and the one at C has two

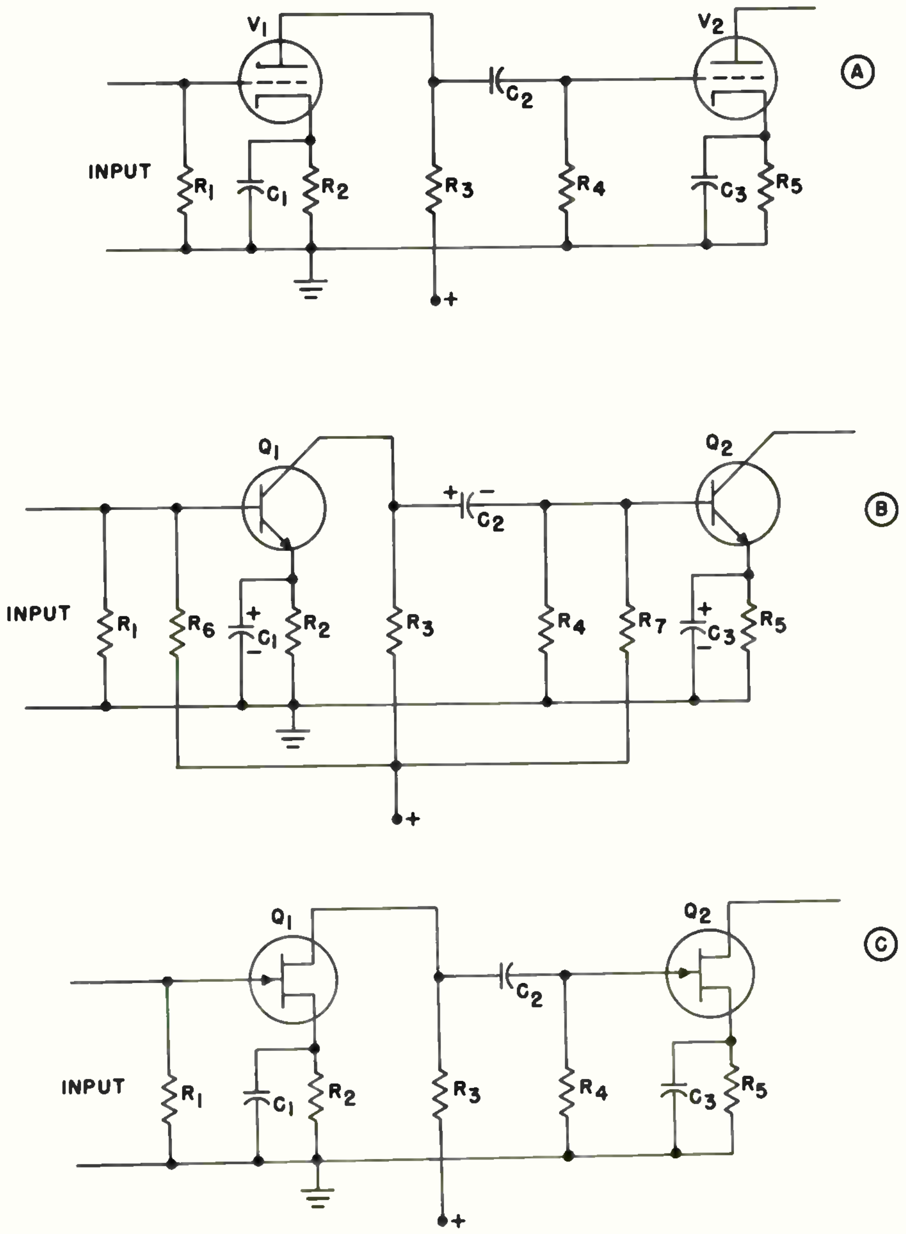


Fig. 16. Basic R-C coupled amplifiers.

field-effect transistors. In each case C_2 is the coupling capacitor between the two stages. The reactance of this capacitor will at some low frequency become equal to the input resistance in the following stage. When this happens, only 70.7% of the voltage developed in the output of the first stage will be fed to the input of the second stage. At lower frequencies, the amount of signal developed by the first stage that reaches the second stage will become increasingly smaller.

In the case of the vacuum-tube amplifier shown in Fig. 16A we can use a large value of grid resistance, R_4, in the input of the second stage. By using as large a resistance as possible in this circuit, we can reduce to a low value the frequency at which the reactance of C_2 becomes equal to the resistance of R_4. In the case of the transistor amplifier shown in Fig. 16B, however, there is not much we can do about increasing the value of the base resistor R_4. The transistor itself draws a certain base current. Thus the transistor itself has a low input resistance. Therefore, this resistance is always in parallel with R_4 and increasing the value of R_4 will have little effect on increasing the actual input resistance of Q_2. We get around the low input resistance of Q_2 by using a high capacity electrolytic capacitor between Q_1 and Q_2 as the coupling capacitor. Since transistors are operated at relatively low voltages we can use a capacitor with quite a high capacity, and still keep the physical size of the capacitor quite small.

In the field-effect transistor circuit shown in Fig. 16C, once again we can increase the value of R_4, the resistance between the gate of Q_2 and the source, to as high a value as possible. Since there is little or no gate current in a field-effect transistor the value of R_4 can be made quite large and increase the input resistance of the second stage. Thus the low-frequency limit of the amplifier can be extended to quite a low value. This is particularly true in the case of the insulated-gate field-effect transistor, since for all practical purposes there is no gate current at all in these transistors.

REDUCING THE EFFECT OF THE COUPLING CAPACITOR

The equivalent circuit of the coupling network used between the two stages in each of the examples shown in Fig. 16 can be redrawn as shown in Fig. 17. In the case of Fig. 16A, using the two vacuum tubes, R_3 is the plate load resistor of V_1, and R_4 is the grid resistor of V_2. In the case of the transistor circuit shown in Fig. 16B, R_3 is the collector resistor of Q_1. R_4 is the resistor connected between the base of Q_2 and ground. In the case of the circuit shown in Fig. 16C where we've used the field-effect transistors, R_3 is the resistor connected between the drain of Q_1 and B+, and R_4 is the resistor connected between the gate of Q_2 and ground.

You can readily see from the circuit shown in Fig. 17 that the output voltage developed at the output of the first stage

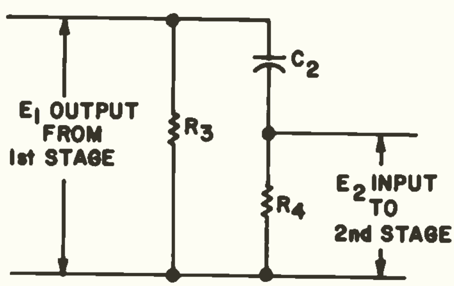


Fig. 17. Equivalent circuit of coupling network at low frequencies.

is fed to the series combination of C_2 and R_4. At high frequencies and middle frequencies, the capacitive reactance of C_2 is so small compared to the resistance of R_4 that practically all of the signal developed at the output of the first stage is fed to the input of the second stage. For all practical purposes, C_2 simply acts like a short circuit.

However, at lower frequencies, the reactance of C_2 begins to become appreciable and it cannot be ignored. At some low frequency, the reactance of C_2 will become equal to the resistance of R_4. Since we have equal capacitive reactance and resistance in this circuit, we will have a current flowing that leads the voltage applied by 45°. Furthermore, since the voltage across C_2 plus the voltage across R_4 must be equal to the voltage across R_3, we have a drop in the voltage across R_4. This means that the voltage applied to the input of the second stage decreases and at the same time we have a phase shift.

The relationship between the output voltage of the first stage which we have labeled E_1, and the input voltage applied to the second stage, which we have labeled E_2, can be seen from the vector diagram of Fig. 18. Notice first, that the voltage E_2 is in-phase with the current I. This is as we might expect because in a resistance, the voltage and current are always in-phase. Therefore the voltage E_2 leads the voltage E_1 by 45°. Notice also that the voltage E_C, which is the voltage across the capacitor, is equal to the voltage E_2. The vector sum of the voltage E_2 and E_C will be equal to the voltage of E_1.

Not only is the drop in voltage across R_4 important, because this will reduce the amplitude of low-frequency signals, but in many applications, the phase shift is just as big a problem. This is particularly true in video amplifiers in television receivers. They must be able to handle very low-frequency signals without any attenuation or phase shift. A phase shift will displace the video information in part of the picture and cause smearing. Therefore it is desirable to keep the drop in low-frequency response and the phase shift as low as possible.

Obviously one of the simplest ways of preventing problems of low frequencies is to use a large value of coupling capacitor. We do this when we design an amplifier to have good low-frequency response. However, in most cases, where you are interested in the low-frequency response of an amplifier, you are also interested in the high-frequency response. Large capacitors also have a higher capacity to ground. Thus there is a limit to how large a coupling capacitor you can use to improve the low-frequency response with-

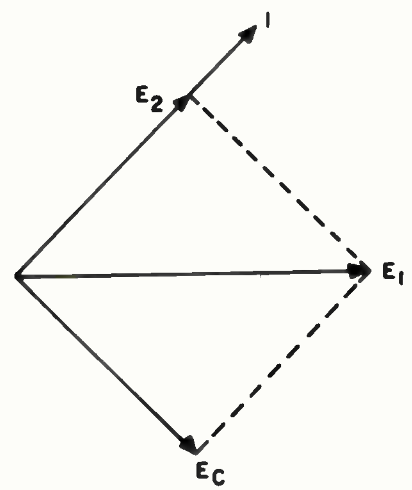


Fig. 18. Vector diagram of voltages at output of first stage and input of second stage.

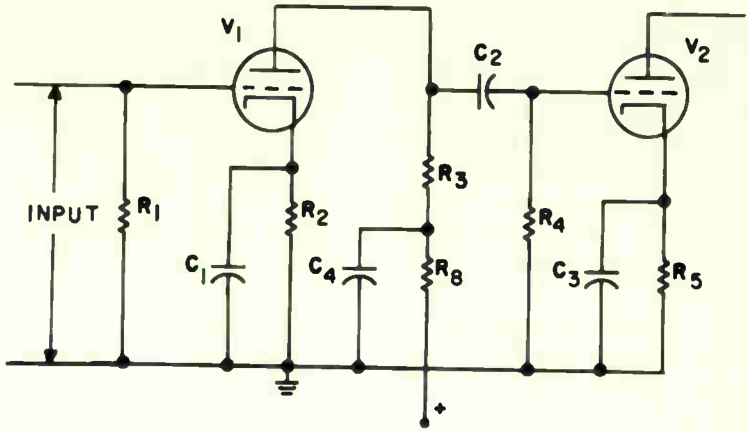


Fig. 19. A two stage amplifier with low-frequency compensation in the plate circuit of V_1.

out running into problems with the high-frequency response. Therefore while we can improve the low-frequency response, to some extent, by using a large coupling capacitor, there is a limit as to how far we can go.

LOW-FREQUENCY COMPENSATION

Regardless of how large a coupling capacitor and of how large a resistance we are able to put into the input of the second stage, the reactance of the capacitor will still increase as the frequency decreases. In some amplifiers where we must amplify very low frequencies of only a few Hertz, with essentially the same gain that we have in the middle-frequency range, we usually have to add something to the circuit to compensate for the fact that the coupling capacitor reactance becomes appreciable at these low frequencies.

In Fig. 19 we have shown how an additional resistor and capacitor can be added in the plate circuit of V_1 in the amplifier shown in Fig. 16A, to compensate for an increase in the reactance of C_2 at low frequencies. In this circuit, the low-frequency compensating network that has been added consists of R_8 and C_4. Let's see how the addition of these two components can improve the gain of the amplifier at low frequencies.

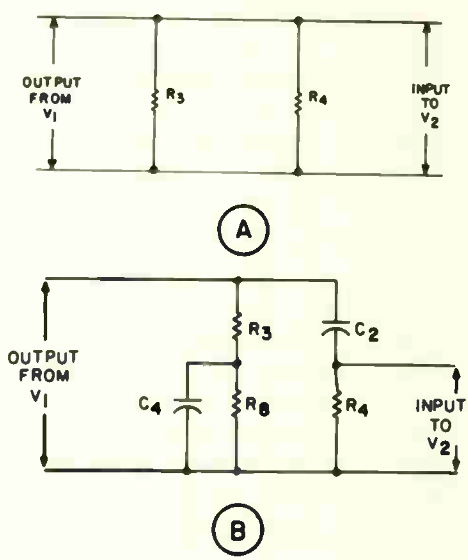


Fig. 20. Equivalent circuit of low-frequency compensation network. "A" is for middle and high frequencies. "B" is for low frequencies.

In Fig. 20A we have shown the equivalent circuit of the coupling network between V_1 and V_2 and the low frequency compensation network at middle and high frequencies. Notice that we have omitted C_2 because in the middle-frequency range, the reactance of C_2 is so low that it can be ignored. Also the value of C_4 is selected so that in the middle and higher frequencies its reactance is very low. It is so low that the end of R_3 that connects to the junction of C_4 and R_8 is in effect connected to ground through C_4. This places R_3 and R_4 directly in parallel. The parallel resistances of these two resistors is the plate load of V_1.

As we pointed out previously, in most amplifiers where we are interested in good low-frequency response we are also interested in good high-frequency response. You will remember that we are going to use a low value plate-load resistor in order to improve the high-frequency response. Therefore, R_3 in Fig. 19 will be a low value resistor. Since at the middle frequencies, R_3 is in effect in parallel with R_4, and R_4 will be many times the value of R_3, the plate load is in effect R_3 alone. For all practical purposes we could omit R_4 in the equivalent circuit shown in Fig. 20A.

At low frequencies, the reactance of C_2 becomes appreciable and it cannot be ignored. At the same time, the reactance of C_4 also becomes appreciable so that it no longer effectively grounds the one end of R_3. Therefore, the plate load for V_1 becomes the combination of R_3 in-series with the parallel combination of R_8 and C_4. In other words, the plate load resistance of V_1 increases. When the plate load of the tube increases, the voltage developed in the plate circuit will increase and this compensates for the loss in voltage across the coupling capacitor C_2.

Not only does the compensating network compensate for the drop in voltage fed to the input of the second stage, but it also improves on the phase shift that occurs at low frequencies. Let's consider how this can happen.

Going back to the equivalent circuit shown in Fig. 20A, remember that the signal current from the vacuum tube develops the voltage across R_3, in other words, across the plate load resistance. This voltage in turn causes a current to flow across the coupling capacitor C_2 and the grid resistor R_4 developing a voltage across R_4. Now consider what happens in the equivalent shown in Fig. 20B. The signal current from the tube flows through the plate load consisting of R_3 in series with the parallel combination of R_8 and C_4. Since the circuit is a capacitive circuit, the voltage developed across this circuit will lag the current. Thus we have a low frequency signal voltage in the plate circuit of V_1 that is lagging the signal current. This lagging signal voltage causes a current to flow through C_2 and R_4. Since the circuit consisting of C_2 and R_4 is capacitive, the current flowing will lead the voltage. By the proper selection of C_4 and R_8 we can cause the phase shift in this network to equal or compensate for the phase shift in the network consisting of C_2 and R_4. Therefore the current flowing through C_2 and R_4 will be in-phase with the signal current supplied by the tube. This means that the signal voltage developed across R_4 will then be in-phase with the signal current supplied by the tube at low frequencies. Since it is already in-phase with the signal current at middle and high frequencies we have compensated for the low-frequency phase shift.

Low-frequency compensation can also be used in circuits where transistors or

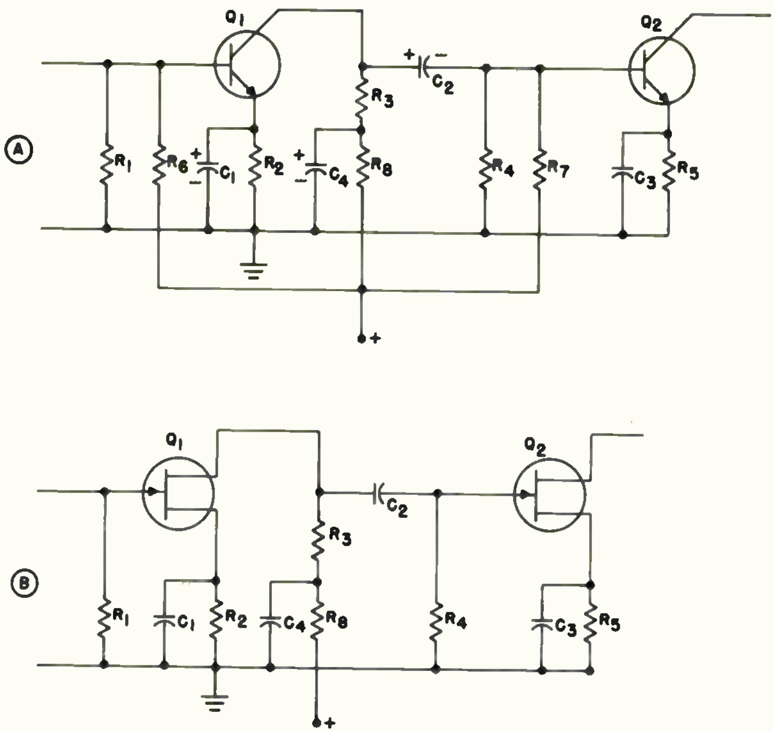


Fig. 21. Low-frequency compensation is shown in a transistor amplifier at A, and in a field-effect transistor amplifier in B.

field-effect transistors have been employed. In the circuit shown in Fig. 21A, we have shown low-frequency compensation in the collector circuit of Q_1. Notice that once again we have added the resistor R_8 in series with the collector-load resistor R_3 and the additional capacitor C_4 in the collector circuit. Low-frequency compensation is not as effective in transistor amplifiers as in vacuum-tube amplifiers.

In the diagram shown in Fig. 21B, we have added low-frequency compensation to the amplifier using field-effect transistors. The additional resistor R_8 is added in the drain circuit of Q_1 along with the additional capacitor C_4. The operation of these low-frequency compensating circuits is exactly the same as in the case of the low-frequency compensating circuit in the vacuum-tube amplifier and is equally effective.

LOW-FREQUENCY DEGENERATION

Going back to the amplifiers shown in Fig. 16, in addition to the coupling capacitor C_2, the bypass capacitors C_1 and C_3 may cause a drop in the low-frequency response. Considering first the vacuum-tube amplifier shown in Fig. 16A, C_1 is the cathode bypass for V_1. The purpose is to maintain the cathode voltage on V_1 constant. At medium and high-frequencies it acts as a bypass capacitor so that the signal current in effect flows through it and R_2 is effectively bypassed. Therefore the voltage across R_2 remains constant. However, at low frequencies, the reactance of C_1 increases. As a result, part of the signal current flows through R_2. This causes a signal voltage to appear at the cathode that is in-phase with the voltage applied to the grid of the tube. This reduces the grid-to-cathode signal voltage so that the output from the stage goes down. This effect can be kept at a minimum by using a large capacitor, usually an electrolytic capacitor, to provide effective bypassing at low frequencies. Another method of eliminating this problem is to eliminate the bypass capacitor altogether. This will reduce the gain of the stage at medium frequencies as well as high frequencies, but the degenerative effect of the unbypassed cathode is constant at all frequencies. Thus the gain of the stage is reduced an equal amount at all frequencies rather than at low frequencies only.

Exactly the same situation exists in the case of the emitter bypass capacitor C_1 in the transistor amplifier circuit shown in Fig. 16B, and the source bypass capacitor C_1 in the field-effect transistor circuit shown in Fig. 16C. These capacitors become degenerative at low frequencies and will reduce the low-frequency response of the stage. By using large electrolytic capacitors the effect can be kept at a minimum so that the response of the amplifier may be satisfactory down to a frequency of only a very few cycles per second. Also the bypass capacitor may be omitted, as in the case of the tube amplifier, introducing a degenerative effect at all frequencies so as to level off the response of the amplifier.

SELF-TEST QUESTIONS

(t) What part is the primary cause of poor low-frequency response in a two-stage amplifier?

(u) In addition to a drop in gain at low frequencies, what other problem is frequently encountered?

(v) What method is used to improve the low-frequency response of amplifiers?

(w) Is low-frequency compensation equally effective in vacuum-tube and transistor-amplifier circuits?

(x) Is low-frequency compensation effective in improving the low-frequency response in an amplifier using field-effect transistors?

(y) What do we mean by cathode degeneration?

Typical Wide-Band Amplifiers

In this section of the lesson we will look at a few typical wide-band amplifiers such as you might encounter in electronic equipment. We will also discuss printed circuit wiring since the trend today in modern electronic equipment is to use printed circuit wiring. This type of wiring is particularly advantageous in wide-band amplifiers because the stray capacities in the circuit can be kept constant from one amplifier to another, and therefore it is quite easy to manufacture amplifiers with almost identical characteristics.

In transistorized equipment, as we mentioned previously, it is not quite as easy to compensate for high-frequency losses as it is in vacuum-tube amplifiers. Therefore we will look at some circuits found only in transistorized equipment that get around some of these disadvantages.

In the preceding section of the lesson we discussed low-frequency compensation. As you will remember, low-frequency compensation is needed primarily because the reactance of the coupling capacitor between stages increases as the frequency goes down. This problem can be overcome by the use of direct coupling; in this type of coupling, the capacitor is omitted entirely and hence any problems created by it are avoided. We will look at direct-coupled amplifiers so you will be familiar with this type of circuit.

There is no doubt that television receivers make more use of wide-band amplifiers than any other electronic equipment. This is simply due to the fact that there are so many television receivers manufactured every year. Therefore we'll start our study of typical wide-band amplifiers with a look at a typical video amplifier which is a wide-band amplifier.

A TYPICAL VIDEO AMPLIFIER

In Fig. 22 we have shown the sche-

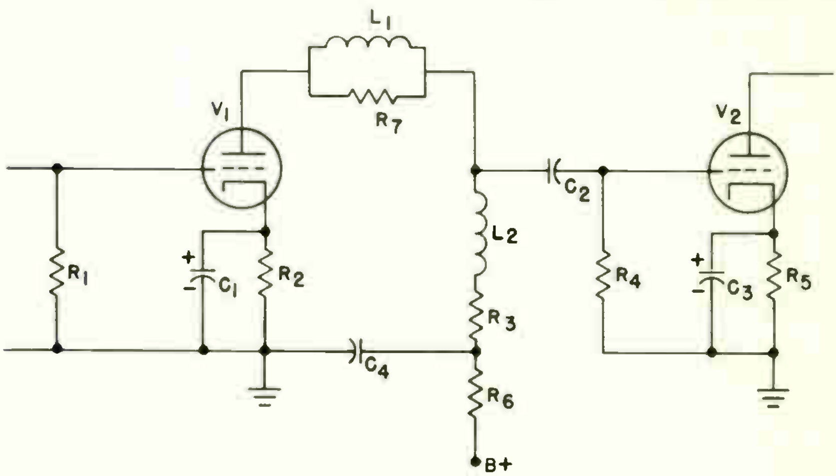


Fig. 22. A typical wide-band amplifier.

matic diagram of a typical wide-band amplifier such as might be used for video amplification in a television receiver. Notice that to improve the high-frequency response of the amplifier both series and shunt peaking have been used. Notice that across the series-peaking coil L_1, we have a shunt resistor. This is to reduce the Q of the coil and prevent oscillation, and at the same time broaden the response of the series-resonant circuit.

L_2 is a shunt-peaking coil and is used to form a parallel-resonant circuit to increase the value of the plate load at high frequencies.

Low-frequency compensation is provided by C_4 and R_6 connected in the plate circuit of V_1. This is similar to the low-frequency compensating network shown earlier. We have shown electrolytic capacitors in the cathode circuits of both stages as the cathode bypass capacitors. You can tell from the schematic when an electrolytic capacitor is used because the polarity of the capacitor is usually indicated. Notice that the polarity signs indicate that the positive side of the capacitor is connected to the cathode of the tube and the other side to ground. Electrons flow from ground through the cathode resistor to the cathode of the tube, and in so doing develop a voltage across the bias resistor having a polarity such that the cathode end is positive. The electrolytic capacitors must be connected with this polarity.

In some wide-band amplifiers, electrolytic cathode bypass capacitors may be shunted by small paper capacitors or ceramic capacitors. You might find a bypass cathode capacitor with a capacity of about 100 mfd shunted by a .001 mfd ceramic capacitor. This is often done because electrolytic capacitors are sometimes rather poor bypass capacitors at high frequencies. Therefore at low frequencies the electrolytic capacitor acts as the bypass capacitor, but at high frequencies where the electrolytic becomes a rather inefficient bypass, the ceramic capacitor has a low enough reactance to bypass the cathode resistor effectively.

TRANSISTOR WIDE-BAND AMPLIFIERS

A transistor video amplifier using PNP transistors is shown in Fig. 23. L_1 is the

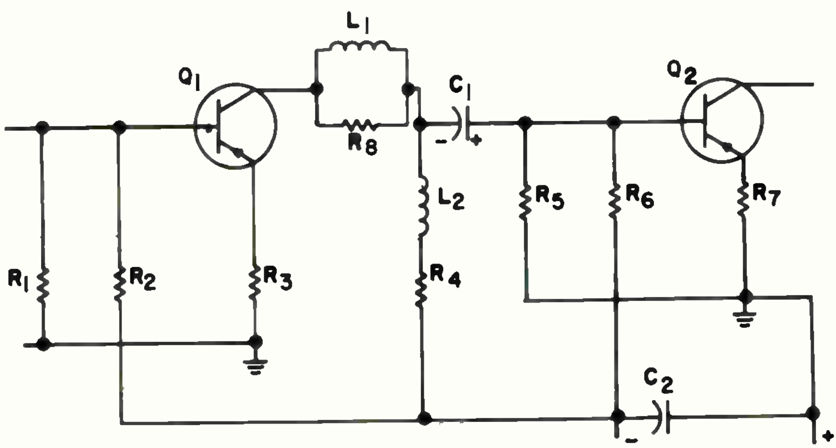


Fig. 23. A transistor wide-band amplifier.

series-peaking coil, and it is loaded by R_8 to reduce the Q of the coil and broaden the response of the circuit. L_2 is the shunt-peaking coil.

The coupling capacitor C_1 is an electrolytic capacitor; a large value of capacitor is used to provide the required low-frequency response. Notice that the emitter resistors R_3 and R_7 are not bypassed. In order to prevent low-frequency degeneration due to the reactance of bypass capacitors across the emitter resistors, the capacitors are simply omitted. As we mentioned previously this introduces degeneration at all frequencies but tends to flatten the gain of the amplifier.

Feedback Circuits. Transistor amplifiers do not lend themselves to improving frequency response by means of peaking coils nearly as well as do vacuum-tube amplifiers. The same is true of improving the low-frequency response; increasing the size of the base resistor of the second stage in a transistor amplifier actually has little effect on the input resistance of the stage. The base current of the transistor primarily determines the input resistance.

In the case of the equivalent coupling circuit for the low-frequency response shown in Fig. 17, in a transistor amplifier R_3 and R_4 are actually shunted by other resistors that should be considered. R_3 is shunted by a resistance which is the output resistance of Q_1, and R_4 is shunted by a resistance which is the input resistance of Q_2. The output resistance of a transistor in a common-emitter circuit is probably around 20,000 ohms whereas the input resistance of a common-emitter circuit is quite low, usually in the vicinity of a few hundred ohms at the most.

The frequency response of transistor amplifiers can frequently be improved by means of feedback circuits. Feedback, of course, must be degenerative feedback and it operates on the principle that as the output signal from the amplifier increases, the amount of feedback must increase. This in turn will tend to reduce the overall gain of the amplifier more at a frequency where the output is higher than it would at another frequency where the output signal is low, and hence the feedback signal is low.

In Fig. 24 we have shown the sche-

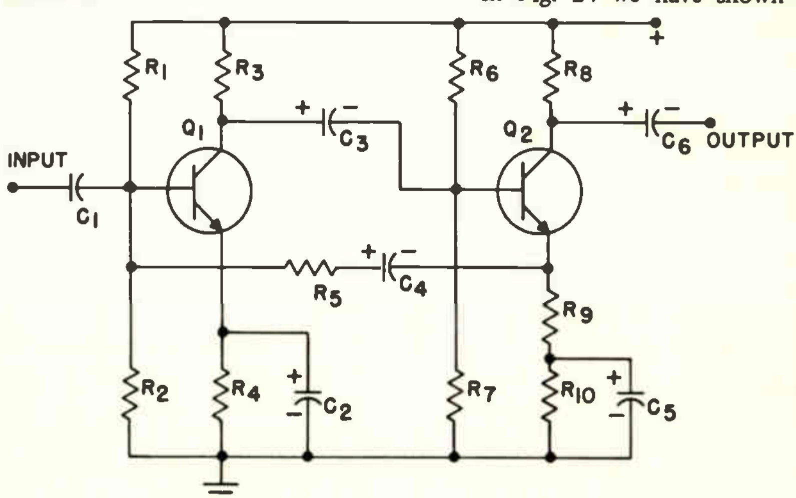


Fig. 24. Improving the frequency response by means of feedback.

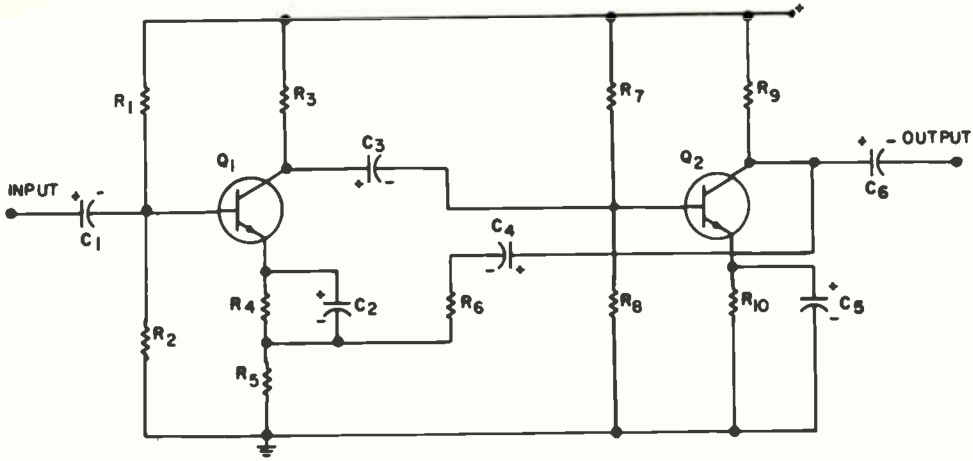


Fig. 25. A second feedback circuit.

matic diagram of a circuit that could be used to improve the frequency response of a two-stage transistor amplifier. The circuit is quite similar to the transistor amplifier shown previously except that the feedback network consists of C_4 and R_5. This provides negative feedback back from the emitter of Q_2 back to the base of Q_1.

Consider what happens as a signal passes through the amplifier. When the signal applied to the base of Q_1 swings in a positive direction, the forward bias across the emitter-base junction of Q_1 will be increased. This will cause the current through the transistor to increase and hence the voltage drop across R_3 will increase. As the voltage drop across R_3 increases, the signal voltage at the collector of Q_1 will swing in a negative direction. Thus for a positive input signal we have an amplified negative signal in the output of Q_1.

The negative signal from Q_1 is fed to the base of Q_2 through the coupling capacitor C_3. The negative-going signal reduces the forward bias across the emitter-base junction of Q_2. Thus the current flowing through Q_2 decreases. When this happens the voltage drop across the emitter resistor R_9 will decrease. The voltage across R_{10} is held essentially constant by C_5. The decrease in signal voltage across R_9 is actually a negative signal. This negative signal is fed through C_4 and R_5 back into the base of Q_1 and subtracts from the original positive signal.

In this circuit, as the gain tends to rise at middle frequencies, the amount of signal fed back from the emitter of Q_2 to the base of Q_1 increases thus tending to reduce the overall gain of the amplifier. On the other hand, if the signal starts to fall off at the high frequencies or at the low frequencies, any amount of feedback will also go down so that the overall gain of the two-stage amplifier will tend to increase.

Another example of feedback which is used to improve the frequency response of a transistor is shown in Fig. 25. Here the signal feedback is taken from the collector of the second transistor Q_2 and fed through a resistor-capacitor network back into the emitter of the first transistor Q_1. To see how this feedback circuit works, let's follow the signal phases through the amplifier.

When a signal drives the base of Q_1 in a positive direction, current through the transistor increases causing a voltage drop across the collector-load resistor R_3 to increase. When this voltage drop increases, the net voltage between the collector of Q_1 and ground swings in a negative direction. Thus in the common-emitter circuit, we always have the situation of a 180° phase shift between the input and output signals. Now the negative-going signal is fed from C_3 to the base of Q_2 and this causes the current through Q_2 to decrease. When the current through Q_2 decreases, the voltage drop across R_9 decreases causing the collector of Q_2 to swing in a positive direction. Once again we have the 180° phase shift in the common-emitter circuit so that we have a positive-going signal produced in the collector circuit of Q_2.

This positive signal is then fed through C_4 and R_6 into the emitter circuit of Q_1. The resistors R_6 and R_5 actually act as a voltage-divider network so that the portion of the signal developed across R_5 will be fed into the emitter circuit of Q_1. Thus when the signal applied to the base of Q_1 swings in a positive direction we also have the emitter of Q_1 swinging in a positive direction due to the feedback signal. This reduces the net base-to-emitter signal voltage so that the increase in current through Q_1 will not be as great as it would be without the feedback.

As in the case of the previous amplifier, when the output from Q_2 begins to increase, the amplitude of the feedback signal will also increase and this in turn tends to reduce the gain of the amplifier to keep the output constant. On the other hand, if the output of Q_2 begins to fall off, then the amount of feedback voltage goes down so that the gain of the amplifier tends to increase and thus level off the overall frequency response of the amplifier.

DIRECT-COUPLED AMPLIFIERS

All of the problems introduced by the coupling capacitor used between two stages can be eliminated by using direct-coupled amplifiers. In the case of a vacuum-tube amplifier, the sole purpose of the coupling capacitor between the two stages is to keep the positive voltage applied to the plate of the tube off the grid of the following tube. We do this so the grid can be operated at dc ground potential and the cathode at a small positive voltage. In this way bias can be obtained for the tube. However, there is no reason why we cannot let the grid operate at a fairly high positive potential and simply operate the cathode at a still higher positive voltage. This will mean that the cathode will be positive with respect to the grid, or in other words, the grid will be negative with respect to the cathode.

A circuit where this is done is shown in Fig. 26. This type of amplifier is called a direct-coupled amplifier. Notice that there is no coupling capacitor used between V_1 and V_2. The circuit of V_1 is more or less conventional, but the circuit of V_2 is somewhat different from the amplifiers you have seen previously. Notice that the grid of V_2 is connected directly to the plate of V_1. It is obvious that with this type of connection we don't have to worry about an increase in coupling capacitor reactance at low frequencies, because there is no coupling capacitor. However, the plate of V_1 must have a positive voltage applied to it, and

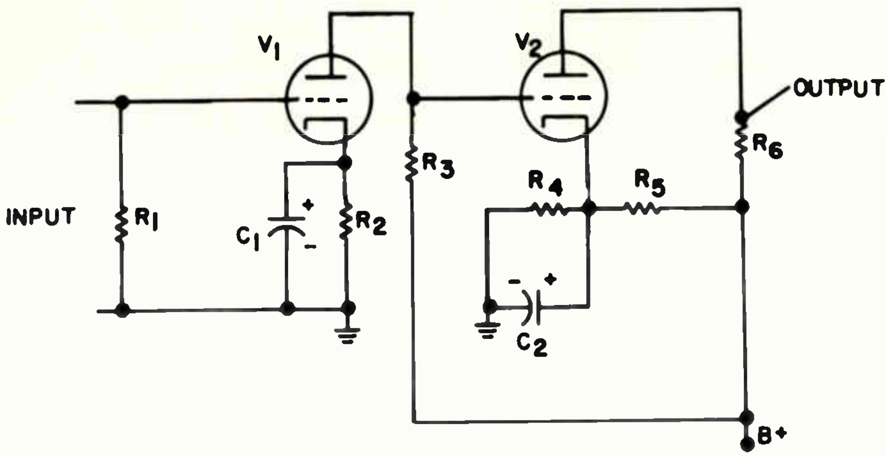


Fig. 26. A direct-coupled vacuum-tube amplifier.

therefore the grid of V_2 will have a positive voltage applied to it. By means of a voltage divider consisting of R_4 and R_5, we have obtained a positive voltage for the cathode of V_2, which is slightly higher than the positive voltage applied to the grid of V_2. C_2 keeps the cathode of V_2 at signal ground potential. Thus we have the signal applied directly between the cathode and grid of V_2. C_2 can be made very large so its reactance is negligible even at very low frequencies. Even if its reactance does start to rise, the value of R_4 is usually somewhat higher than the cathode resistor of a conventional R-C coupled stage. Thus the capacitor does not have to bypass such a low value resistor and more effective bypassing can be obtained.

One disadvantage of this type of amplifier is that aging of one tube appreciably affects the operation of the other. For example, if the emission of V_1 drops, the plate current drawn by this tube will drop. This means that the voltage drop across R_3 will decrease, and therefore the plate voltage on V_1 will increase. This increase in voltage might be enough to swing the grid of V_2 positive with respect to the cathode. Needless to say, when this happens, V_2 starts drawing a higher than normal current. Fortunately, when the tube starts drawing a higher than normal current, the voltage drop across R_4 increases and this tends to compensate for the increase in grid voltage to some extent. In some direct-coupled amplifiers you will find that the plate-load resistor R_3 is returned to a lower voltage than the plate-load resistor R_6. This makes it possible to use a lower value resistor for R_3 and thus tends to give better high-frequency response. You will remember that one of the first steps we did in order to get good high-frequency response from an amplifier was reduce the value of the plate-load resistor. This doesn't actually bring up the response but it does reduce the middle and lower frequencies so that we get a constant gain over a wider frequency range.

TRANSISTOR DIRECT-COUPLED AMPLIFIERS

Transistors lend themselves very well to direct coupling. An example of several transistors used in a direct-coupled circuit

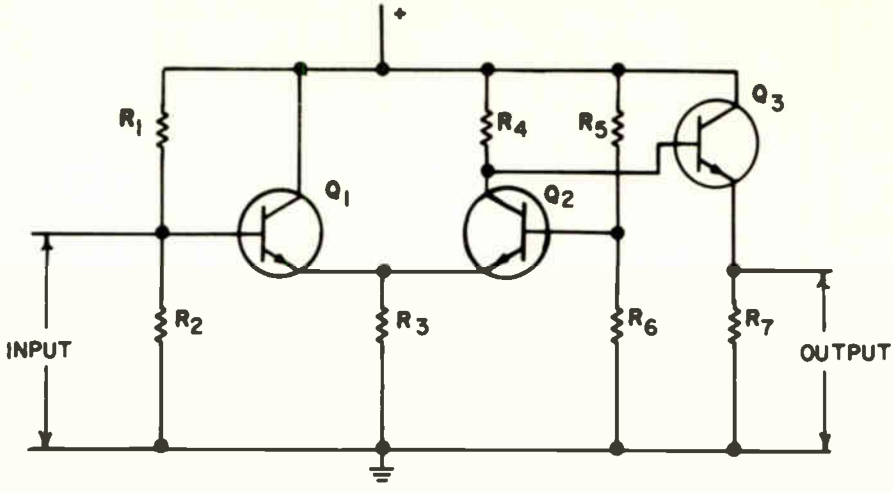


Fig. 27. Transistor direct-coupled stages.

is shown in Fig. 27. The resistors R_1 and R_2 provide a forward bias across the emitter-base junction of the NPN transistor Q_1. Forward bias across the emitter-base junction of Q_2 is provided by R_5 and R_6.

To see how the amplifier works, let's consider what happens when an input signal is applied between the base of Q_1 and ground. When the signal swings the base of Q_1 in a positive direction, current through Q_1 will increase. This causes the voltage drop across R_3 to increase. The increased voltage drop across R_3 increases the emitter voltage on Q_2 thus reducing the forward bias across the emitter-base junction of Q_2. Hence the current through this transistor goes down, and the voltage drop across R_4 will decrease. Therefore the collector of Q_2 will swing in a positive direction.

The collector of Q_2 is connected directly to the base of Q_3 which is used in an emitter-follower circuit. The dc voltage developed across R_7 is relatively low, so the output can be fed directly to the base of another amplifier configuration like Q_1 and Q_2. Of course, the varying signal fed to the base of Q_3 will cause the voltage across R_7 to vary.

The combination of transistors Q_1 and Q_2 is called a differential amplifier. This type of amplifier is widely used in integrated circuits that are used in some television receivers. The purpose of using the emitter follower after Q_2 is to reduce the dc voltage level. If we simply fed the base of a second differential amplifier pair from the collector of Q_2, the base of the following transistor would be at a higher potential than the base of Q_1. Thus for each two-stage differential amplifier we went through, the voltage would be gradually increasing. However, by using an emitter follower, the dc voltage across the emitter-resistor R_7 is quite low so that we can get back down to a low base voltage such as we had on Q_1 originally. By using a two-stage differential amplifier followed it by an emitter follower we can cascade groups of these stages any number of times to get the gain we need, while the operating voltage requirements are quite low.

PRINTED CIRCUIT BOARDS

One of the most common problems encountered in the early days of television, and when manufacturers first began to build wide-band amplifiers for other uses, was getting consistent results. In building a wide-band amplifier, stray capacities have an appreciable effect on the high-frequency response of the amplifier. One amplifier might have good frequency response up to 5 MHz and the next one built by the manufacturer, with the same parts and tubes, might have as good a response to only 4 MHz (or less). Often this is due simply to parts being put in slightly different positions and wires that were routed slightly different by the person who assembled the equipment.

Problems of this type have been almost completely eliminated by the use of printed circuit boards in the assembly of wide-band amplifiers. In the printed circuit board the wiring is actually on a phenolic or glass epoxy-type base. The wiring consists of copper strips that are glued to the base.

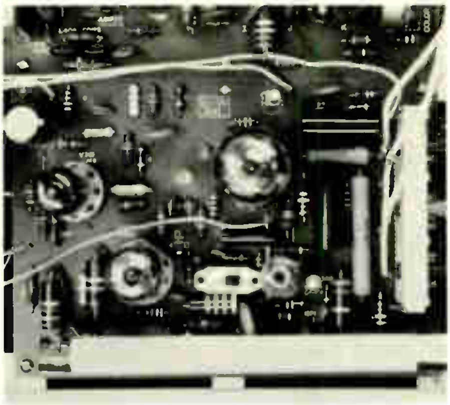
In manufacturing a printed circuit board the manufacturer starts off with a board that may have been made out of a phenolic material that is 1/16th of an inch thick. This material has a thin sheet of copper firmly glued to one side. The basic boards are available in sheets that are 4′ wide by 8′ long. The manufacturer cuts the desired size board from sheets of this type. In manufacturing the board for use in wide-band amplifiers, the wiring on the board is etched. To do this we simply draw the circuit and then transfer it photographically to the copper circuit board. The board is then placed in an etching solution and the undesired copper is etched away leaving only the copper desired to make the connections between the various parts on the board. The parts needed are then mounted on the other side of the board and connected to the copper wiring by inserting the part leads through holes that have been drilled or punched in the board. The leads to the various parts are then soldered to the copper wiring.

The advantage of an etched circuit board, of course, is that all parts fall into exactly the same place, the wiring falls in exactly the same place and hence distributed and stray capacities remain essentially constant from one board to the next. Thus assuming we hold the parts tolerances to a reasonable value, we can expect the frequency-response and gain of one amplifier to be essentially the same as the next one we manufacture.

In Fig. 28 we have shown a photograph of the circuit board containing the video amplifier of the Conar Model 600 Color Television Receiver. Notice that the tubes and parts are all mounted on one side of the circuit board and the wiring which interconnects all these components is in the form of copper on the other side of the board. The use of the printed circuit board ensures that video amplifiers of the various sets will have essentially the same frequency response.

One of the disadvantages of circuit boards is that they sometimes develop cracks. This may be due to rough handling of the equipment in which they are used. Sometimes the crack is so thin that you can't see it; sometimes we refer to this as a hairline crack. The crack in the circuitry may cause the equipment to perform intermittently.

Hairline cracks are quite difficult to locate, but they usually can be located by putting a little pressure on the board in various spots. You can do this with some



A

B

Fig. 28. Video amplifier on an etched circuit board. This type of wiring is referred to as printed wiring.

insulated tool; you will find that as you push in a certain spot the intermittent can be made to occur. Look for a hairline break in the copper wiring somewhere near this point. If you cannot find it, simply flow solder over the copper conductors in the area where you think the break is located and then try flexing the board again and see if the trouble is eliminated. The solder will usually bridge right across the crack and eliminate the intermittent.

Sometimes in repairing equipment in which a printed circuit board is used you'll accidentally pull the copper loose from the board. The copper is only glued to the board and can be easily knocked off by applying too much heat when you are replacing a part. In this case, you can effect a repair simply by taking a piece of wire and soldering the wire in place to complete the circuit between the two points from which the copper has been accidentally removed. Use a short direct piece of wire so you do not upset the performance of the board.

Printed circuit boards have added much insofar as obtaining consistent performance from wide-band amplifiers is concerned. They are not particularly difficult to work on as long as you make repairs quickly and avoid applying excessive heat or force to the board. Some servicemen do not like to work on printed circuit boards because it is somewhat more difficult to trace out the circuit than with the older style of hand wiring, but with a little practice you can learn to do this. Since most modern radio and TV receivers use printed circuitry it is important that you learn to work with this type of equipment.

SELF-TEST QUESTIONS

(z) What system is frequently used in transistor amplifiers to improve the overall frequency response of the amplifier?

(aa) How can you repair a "hairline" crack in a printed circuit board?

(ab) How can you repair a printed circuit board in which a piece of the copper has been accidentally removed?

(ac) What is the advantage of using an emitter follower after a differential amplifier such as shown in Fig. 27?

(ad) In the direct-coupled amplifier shown in Fig. 26, how is the high positive voltage on the grid of V_2 overcome?

(ae) What type of high-frequency compensation is used in the circuit shown in Fig. 23?

(af) What components form the low-frequency compensating network in the amplifier shown in Fig. 22?

LOOKING AHEAD

You can look forward to servicing many amplifiers in your career as an electronics technician. You will find amplifiers used in all types of electronic equipment. Wide-band amplifiers will be found in many different applications. Make sure that you understand wide-band amplifiers before leaving this lesson and going on to the next. A small amount of additional time spent on this lesson may save you a great deal of time later when you start working on equipment of this type.

Answers to Self-Test Questions

(a) A common logarithm is a number which tells us the power to which 10 must be raised in order to equal a given number.

(b) The characteristic.

(c) The mantissa.

(d) Two. The characteristic will be one less than the number of digits in the number. Since all numbers between 100 and 999 have three digits, the characteristic will be 2.

(e) The logarithm of 7 is .8451. The characteristic is 0, so it is simply omitted.

(f) The logarithm of 700 is 2.8451. The mantissa is the same as for the logarithm of 7 or for the logarithm of 70 for that matter. The characteristic, which is 2, indicates that the number is somewhere between 100 and 999.

(g) The logarithm of 41.7 is 1.6201. We find this logarithm by locating 41 in the N column of the log table and then moving over to the column under 7 where we see that the mantissa is .6201. 41 has two digits in it so we know that the characteristic must be 1 and therefore the complete logarithm is 1.6201.

(h) The bel is a power ratio. It is equal to the logarithm of the ratio of one power to another power.

(i) The decibel is one tenth of a bel. A decibel is a power ratio and it is equal to ten times the logarithm of one power divided by another. It is expressed by the formula:

$$\text{db} = 10 \log \frac{P_1}{P_2}$$

(j) 7.782 db. To find the change in db we let P_1 equal 222 watts and P_2 equal 37 watts. Now we substitute these values in the decibel power formula and we get:

$$\text{db} = 10 \log \frac{222}{37}$$

dividing 37 into 222 gives us 6 so the change in db equals

$$\text{db} = 10 \log 6 = 10 \times .7782 = 7.782$$

(k) 6 db. To find the answer we use the voltage formula:

$$\text{db} = 20 \log \frac{E_1}{E_2}$$

and we substitute 150 volts for E_1 and 75 volts for E_2 so that we have:

$$\begin{aligned} \text{db} &= 20 \log \frac{150}{75} \\ &= 20 \log 2 \\ &= 20 \times .301 \\ &= 6.0 \text{ db} \end{aligned}$$

(l) To increase the frequency at which the reactance of the shunt capacities becomes equal to the resistance of the plate-load resistor.

(m) If we use too large a grid resistor in a vacuum-tube amplifier, electrons accidentally striking the grid and flowing through the resistor may build up an appreciable bias in the grid of the tube. Another possibility is that the tube may have a small amount of gas. The gas molecules striking the grid will cause a current to flow through the grid resistor that will place a positive bias on the grid of the tube. If we use too large a grid resistor this positive bias may be high enough to cause an excessive current to flow through the tube and destroy the tube.

(n) At high frequencies L_1 resonates with the capacity in the circuit and hence tends to increase the value of the plate load of V_1. Thus the effect of the shunting capacity is reduced.

(o) L_1 is a series-peaking coil. It separates the shunt capacity in the output of V_1 from the shunt capacity in the input of V_2. It forms a series-resonant circuit with the capacity in the input of V_2 and hence builds up the voltage applied to the grid of V_2 at high frequencies.

(p) By series-shunt peaking we mean a combination of shunt peaking such as shown in Fig. 5 and series peaking such as shown in Fig. 8. A series-shunt compensated amplifier is shown in Fig. 10.

(q) No. The low output resistance of the first stage along with the low input resistance of the second stage reduces the effective value of the collector-load resistor in the first amplifier stage. Therefore it takes a larger shunt capacity to produce a capacitive reactance equal to the resistance of the emitter load. As a result, the response of a transistor amplifier frequently starts to fall off at a somewhat higher frequency than the response of a vacuum-tube amplitier.

(r) Yes, the time that it takes the electrons to cross the base of a transistor may become equal to the period of one half cycle at some high frequency. When this happens, instead of getting a 180° phase shift in the common-emitter transistor circuit, we get a 360° phase shift so that the signal in the emitter circuit will be in-phase with the signal in the base circuit. This will cause positive feedback through the transistor and cause the transistor to go into self oscillation.

(s) More like those encountered in vacuum tubes. Because of the high input and output resistances of the field-effect transistor, the field-effect transistor resembles a vacuum tube more closely in performance than it does a transistor. Thus series and shunt peaking are quite effective in improving the high-frequency response of an amplifier using field-effect transistors.

(t) The reactance of the coupling capacitor used between the two stages becomes too high.

(u) Phase shift.

(v) Low-frequency compensation. This usually consists of adding a resistance and a capacitance in the output circuit of the first stage. By selecting the correct value of resistance and capacitance, both the gain and the phase shift can be improved.

(w) No, it is more effective in vacuum-tube circuits, but can be used to some extent in transistor circuits.

(x) Yes. Field-effect transistors have characteristics similar to vacuum tubes and low-frequency compensation is equally effective in the two.

(y) By cathode degeneration we are referring to the reactance of the cathode bypass capacitor becoming so high that the capacitor is no longer an effective bypass. This occurs at very low frequencies and causes poor low-frequency response. It can be overcome by omitting the cathode bypass capacitor so that the degeneration becomes constant at all frequencies. This levels off the gain of the amplifier.

(z) Negative feedback. In a two-stage amplifier, a signal is taken from the second stage and fed back to the first stage to detract from the original signal applied to the first stage. Thus if the output signal of an amplifier tends to increase, the feedback increases reducing the overall gain of the amplifier. On the other hand, if the output tends to fall off, the feedback signal decreases, allowing the amplifier gain to increase.

(aa) By applying a small amount of pressure to the circuit board you can frequently isolate the area in which the crack exists. Then by flowing solder over the copper in that area you can frequently bridge the gap.

(ab) You can bridge the copper by means of a piece of hookup wire. The two ends of the hookup wire are simply soldered in place so that the wire takes the place of the copper that has been accidentally removed.

(ac) It reduces the dc voltage level in the circuit so that the next differential amplifier pair can be operated at the same dc potential as the first. This keeps the power supply requirements quite modest.

(ad) The bias network consisting of R_4 and R_5 places the cathode of V_2 at a higher positive potential than the grid. Thus even though the grid is positive with respect to ground, it is still negative with respect to the cathode.

(ae) Both series and shunt peaking are used.

(af) R_6 and C_4.

Lesson Questions

Be sure to number your Answer Sheet B204.

Place your Student Number on every Answer Sheet.

Most students want to know their grade as soon as possible so they mail their set of answers immediately. Others, knowing they will finish the next lesson within a few days, send in two sets of answers at a time. Either practice is acceptable to us. However, don't hold your answers too long; you may lose them. Don't hold answers to send in more than two sets at a time, or you may run out of lessons before new ones can reach you.

1. What is the logarithm of 2000?

2. Suppose the power output from power amplifier A is increased from 100 to 200 watts, and the power output from power amplifier B is increased from 1 to 5 watts. Which change is the larger change in decibels?

3. Why can the coupling capacitor used between two resistance-capacitance coupled stages be ignored at the middle and high frequencies?

4. What limits the high-frequency response of an R-C coupled amplifier?

5. What effect does reducing the size of the plate load resistor in an amplifier to extend the high-frequency response have on the gain at the middle and low frequencies?

6. Name the two types of peaking used to extend the high-frequency response of an amplifier.

7. Why can we not increase the size of the base resistor R_4 in a transistor amplifier such as shown in Fig. 16B to improve the low-frequency response of the two stage amplifier?

8. Why is an electrolytic capacitor used as a cathode bypass or an emitter bypass sometimes shunted by a paper or a ceramic capacitor?

9. What method can be used other than peaking coils and low-frequency compensation to improve the frequency response of an amplifier?

10. In the direct coupled transistor amplifier circuit shown in Fig. 27, what primary purpose does Q_3 serve?

REVERE OUR LAWS

Let every American, every lover of liberty, every well-wisher to his posterity swear by the blood of the Revolution never to violate in the least particular the laws of the country - Let every man remember that to violate the laws is to trample on the blood of his father, and to tear the charter of his own and his children's liberty.

Let reverence for the laws be breathed by every American mother to the lisping babe that prattles on her lap; let it be taught in schools, in seminaries, and in colleges; let it be written in primers, spelling books and in almanacs; let it be preached from the pulpit, proclaimed in legislative halls and enforced in courts of justice. And in short, let it become the political religion of the nation; and let the old and the young, the rich and the poor, the grave and the gay of all sexes and tongues and colors and conditions, sacrifice unceasingly upon its altars.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-217